

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 16:09:21 : Search time 71.77 Seconds  
(Without alignments)  
253.023 Million cell updates/sec

Title: US-09-052-089a-1

Sequence: 1 MPRLCTICSDFFDHSRDV.....VRKTVPSLFOAKLDTFLMS 469

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query %	Match length	ID	Description
1	192.5	8.1	1938	1 MYHD_HUMAN	Q9UKX3 homo sapien
2	187.5	7.9	2663	1 CENE_HUMAN	Q02224 homo sapien
3	183.5	7.7	1957	1 YD86_SCHPO	Q10411 schizosacch
4	182.5	7.7	1290	1 XPC_XENLA	P50332 xenopus lae
5	178	7.5	962	1 VDP_HUMAN	P06763 homo sapien
6	175	7.4	1509	1 MYSN_ACACA	P05659 acanthameb
7	173.5	7.3	1935	1 MYSS_CTPCA	Q90339 cyprinus ca
8	172	7.2	1938	1 MYH4_RABIT	Q28641 oryctolagus
9	172	7.2	2116	1 MYSD_DICDI	P08739 dictyosteli
10	170	7.1	978	1 RASO_AOUAE	Q67134 aquifex aeo
11	169.5	7.1	1744	1 TANA_XENLA	Q01550 xenopus lae
12	169.5	7.1	1940	1 MYH3_RAT	P12847 rattus norv
13	169	7.1	1935	1 MYH7_RAT	P02564 rattus norv
14	168	7.0	941	1 VDP_MOUSE	Q92150 mus musculu
15	168	7.0	1325	1 G160_MOUSE	P55937 mus musculu
16	168	7.0	1941	1 MYH2_HUMAN	Q9UKX2 homo sapien
17	167	7.0	879	1 MYSP_ONCVO	Q02171 onchocerca
18	167	7.0	880	1 MYSP_BRUMA	Q01202 brugia mala
19	167	7.0	2411	1 MYSA_DROME	P05661 drosophila
20	166	7.0	388	1 PAM_STRPY	P49034 streptococ
21	166	7.0	959	1 VDP_RAT	P15422 rattus norv
22	166	7.0	1935	1 MYH7_PIG	P79233 sus scrofa
23	166	7.0	1939	1 MYH1_HUMAN	P12882 homo sapien
24	166	7.0	1960	1 MYH9_HUMAN	P35579 homo sapien
25	166	7.0	2704	1 BPAL_HUMAN	Q03001 homo sapien
26	165.5	6.9	1790	1 USOL_YEAST	P25366 saccharomyc
27	165	6.9	848	1 MYSP_DIRIM	P43332 dirofilaria
28	165	6.9	886	1 RASO_ARCFU	Q29230 archaeoglob
29	164.5	6.9	880	1 RASO_PVRAB	Q09268 pyrococcus
30	163.5	6.9	483	1 M6_STRPY	P08089 streptococ
31	163.5	6.9	1453	1 Y373_BOVIN	Q9LUA3 bos taurus
32	163.5	6.9	1639	1 LMGI_DROME	P15215 drosophila
33	163.5	6.9	1727	1 ALMI_SCHPO	Q9UKX5 schizosacch

34	163.5	6.9	1976	1 MYH4_HUMAN	P35580 homo sapien
35	163.5	6.9	2871	1 DESP_HUMAN	P15924 homo sapien
36	163	6.8	1935	1 MYH7_HUMAN	P12883 homo sapien
37	163	6.8	1938	1 MYSS_CHICK	P13538 gallus gall
38	162.5	6.8	1232	1 KF4A_HUMAN	Q95239 homo sapien
39	162.5	6.8	1938	1 MYA_AEOIR	P24733 aequipecten
40	162.5	6.8	3911	1 AKAA_HUMAN	Q99996 h a-kinase
41	162	6.8	963	1 KINH_HUMAN	P33176 homo sapien
42	162	6.8	1966	1 MYSB_CAEEL	P02566 caenorhabdi
43	161.5	6.8	1939	1 MYH4_HUMAN	Q9Y623 homo sapien
44	161.5	6.8	1969	1 MYSA_CAEEL	P12844 caenorhabdi
45	161.5	6.8	1972	1 MYHB_MOUSE	Q08638 mus musculu

## ALIGNMENTS

```

RESULT 1
ID MYHD_HUMAN STANDARD: PRT: 1938 AA.
AC Q9UKX3: Q95252;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, extraocular (MYHC-ec).
GN MYH13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99318869; PubMed=10388558;
RA Weiss A., Schiaffino S., Lelwand L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: implications for functional diversity.";
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1917-1938 FROM N.A.
RC TISSUE=Extraocular muscle;
RX MEDLINE=99026150; PubMed=9806854;
RA Winters L.M., Briggs M.M., Schachar F.;
RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
RT the cluster of fast and developmental myosin genes on chromosome 17.";
RL Genomics 54:186-189(1998).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
DR EMBL: AF111782; AAC29948.1; -
DR EMBL: AF075248; AAC83241.1; -

```

DR HSSP; P08799; 1MND.  
 DR MIM; 603487; -.  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 2.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PD00193; MYOSINHEAVY.  
 DR PRODOM: PD00035; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM Calmodulin-binding; ATP-binding; Methylation; Alkylation;  
 KM Multigene family.  
 KM Multigene family.  
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.  
 FT DOMAIN 785 814 IQ.  
 FT DOMAIN 843 1938 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT DOMAIN 659 681 ACTIN-BINDING (BY SIMILARITY).  
 FT DOMAIN 761 775 ACTIN-BINDING (TR1-) (POTENTIAL).  
 FT MOD\_RES 130 130 METHYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 699 699 ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 709 709 ALKYLATION (SH-2) (POTENTIAL).  
 SO SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CDS CRC64;

Query Match 8.1%; Score 192.5; DB 1; Length 1938;  
 Best Local Similarity 20.9%; Pred. No. 0.018;  
 Matches 94; Conservative 74; Mismatches 161; Indels 121; Gaps 16;

QY 3 IRACTICSDFFD--SRDVAIHGHTFHQCICQSFETAPSTCQCRCRYQVKRTIIN 60  
 DB 1249 IERTCTVEQFSEIKAKDEQQTQLIHLDNQG-----KARLTQNGELSH 1293  
 QY 61 KLFEDLAQEEENVL-----DREFIKNELDNVRAQLSQDKREK-----RDSQVITDILR 108  
 DB 1294 RV-----EKRESLSQLTGKSKQALTOLEELKRMEEETAKNMAHALOSSRDCCDLR 1348  
 QY 109 DYLEENATVYSLQALGKAE-----MICSTLKKOMRYLEQOODETKQAQ 153  
 DB 1349 EOYEEOEAKAEIQRALSKANSEVAQWKTYETDAIQRTLEBAKKKLAQRIQEAEREK 1408  
 QY 154 EEEGRLLSKKMTMOEIOILLLOSLPEVEEMIRMGVGSQSAVEQLAVVCSLAKREYENLK 212  
 DB 1409 ETA---NSKCAISLEKTKORLQG---EVEDLMRDL-----ERSHTACATLDKKORNFDK 1455  
 QY 213 -----BARKASGEVADK---LRKDLFFSSRSKLTQTVYSELDAQKLELSAQKDLQS 259  
 DB 1456 VLAEMWKQKLDQESQELBAQKESRSLSTELFKMNAVEEYVDQLETLRREKKNQGEISD 1515  
 QY 260 ADEEINLKKKLMLQET-----LNLPPYASE 286  
 DB 1516 LTRQIAETGKNGLEAEKTKLVLEQKESDLOVALEEVGSLSEHRSKILRYOLELSQYKSE 1575  
 QY 287 TVDRLVESAPAEVNVNKLRRPSRDDIDNATPDVD---TPRPPSSSGHGYYETLCL 342  
 DB 1576 -LDRKYKEDEEIEF---QLKRNQRAAEALQSVDAIEIRSNDAIRLKKKMEGDLNMEI 1631  
 QY 343 EKSHSPIQDVPK---KICKGPRKESQSL 368  
 DB 1632 QLGSHSNQMAETQRHLRTVQGLKDSQLHL 1661

RESULT 2  
 CENE\_HUMAN STANDARD; PRT; 2663 AA.  
 AC 002224;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update).  
 DE Centromeric protein E (CENP-E protein).  
 GN CENPE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93024922; PubMed=1406971;  
 RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;  
 RT "CENP-E is a putative kinetochore motor that accumulates just before  
 RT mitosis.";  
 RL Nature 359:536-539(1992).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95196755; PubMed=7889940;  
 RA Thirmer D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;  
 RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed  
 RT microtubule motor.";  
 RL EMBO J. 14:918-926(1995).  
 RN [3]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98437347; PubMed=9763420;  
 RA Chan G.K.T., Schaar B.T., Yen T.J.;  
 RT "Characterization of the kinetochore binding domain of CENP-E reveals  
 RT interactions with the kinetochore proteins CENP-F and hNubrl.";  
 RL J. Cell Biol. 143:49-63(1998).  
 CC -I- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE  
 CC KINETOCHORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE  
 CC OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT  
 CC AND/OR SPINDLE ELONGATION.  
 CC -I- SUBUNIT: INTERACTS WITH CENP-F AND BUBR1 KINASE.  
 CC -I- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCHORES DURING  
 CC CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE. AND IS  
 CC QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.  
 CC -I- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z15005; CA78727.1; -.  
 DR PIR: S28261; S28261.  
 DR HSSP: P17119; 3KAR.  
 DR MIM: 117143; -.  
 DR InterPro: IPR001752; kinesin.  
 DR Pfam: PF00225; kinesin\_1.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR SMART: SM00129; KISC; 1.  
 DR PROSITE: PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE: PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KM Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;  
 KM Cell cycle; Centromere.  
 FT DOMAIN 1 335 KINESIN-MOTOR.  
 FT DOMAIN 336 2471 COILED COIL (POTENTIAL).  
 FT NP\_BIND 2472 2663 GLOBULAR (POTENTIAL).  
 FT NP\_BIND 86 93 ATP (BY SIMILARITY).  
 SO SEQUENCE 2663 AA; 312087 MW; CBF313880C8C8CB8 CRC64;

Query Match 7.9%; Score 187.5; DB 1; Length 2663;  
 Best Local Similarity 25.1%; Pred. No. 0.045;  
 Matches 93; Conservative 66; Mismatches 133; Indels 79; Gaps 18;

QY 55 KRTINKLFEDLAQEEENVLDREFIKNELDNVRAQLSQDKREK-RDSQVITDILRPTLEE 113  
 DB 1632 KMTAVN-----ETQKCKEIEHLKEQFETQKLNLENITENIRLITQI---LHENLEE 1680





```

RL J. Cell Biol. 105:913-925(1987).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEAVYMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -1- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMAIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00624; CAA68663.1; -.
DR PIR: A27224; A27224.
DR HSSP: P08799; IAMD.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR PRODOM: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW Methylation; Alkylation; Phosphorylation; Multigene family.
FT DOMAIN 1 789 MYOSIN HEAD-LIKE.
FT DOMAIN 790 819 IQ.
FT DOMAIN 848 1509 COILED COIL (POTENTIAL).
FT DOMAIN 848 1226 ALPHA-HELICAL TAILPIECE (S2).
FT DOMAIN 1227 1252 HINGE.
FT DOMAIN 1253 1509 LIGHT MEROMYOSIN (LMM).
FT DOMAIN 1253 1482 ALPHA-HELICAL TAILPIECE (LMM).
FT DOMAIN 1483 1509 NONHELICAL TAILPIECE.
FT NP_BIND 182 189 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 766 780 ACTIN-BINDING.
FT MOD_RES 133 133 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 1489 1489 PHOSPHORYLATION.
FT MOD_RES 1494 1494 PHOSPHORYLATION.
FT MOD_RES 1499 1499 PHOSPHORYLATION.
SQ SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

```

```

Query Match 7.4%; Score 176; DB 1; Length 1509;
Best Local Similarity 23.4%; Pred. No. 0.088;
Matches 97; Conservative 74; Mismatches 156; Indels 88; Gaps 13;

```

60 NKLTFDLAQQEENVLDRELFKLNELDNVRAOISOKD-----KEKRSD 101

894 DKLEKDLALALKLIDEGEKADLEEDNALQKVAAGLEELQERTSANDILQQRKLE 953

102 VIIDTLRDLTEERNATVVSLOALGKAE-----MLCSTLKQMKMYLEQOQD 147

954 AEKGLKASLDEEERNRAALQDAKTVSESENNELQDKEDAAHADSILKKKEEDSLRLR 1013

148 ETQKQAEAGR---LRSKMKTMQEIILLLOSLPEV-EEIMRDVGQSAVEOLAYVCV 202

1014 ETNKLADAENISEETLSKLTNTERGADVNEIDVTATKLOLEKTKSLEBELAQTRA 1073

203 SLKREYENLEKARSGEVDKLRKLDPFSSRSKIQOTYSELDQAKLEKSAO---KDLQ- 258

1074 QLEBE-----KSGKRAASAKKQLOGLQEDARSEVDSLKSLSAEKSILAKRADOONRLDE 1129

259 -----SADKEINSLKKLITMLQETL-----NLPPVASETVDRVLSPAPE- 300

1130 QLEDEFTVANVDKQKALEKITELBDQYALDQGNNAQAQKLTQVDEFTKRLEE 1189

301 -----VNLKLRRPSFRDDIDNATFDVDTTPPARSSQHGYYELCLCKSHSPLODVPK 355

1190 AEASARLKERKKAALDEV---ADLTADLDAERDSGAQORRKLMTRISELQSELENAVK- 1245

356 ICKPFRRESQSLISGGQSCAGPEDEIVGAPPIFRNAILQOKOPKRRPSSSCSK 410

1246 -TGGASSEEVKRLGE--LERLEBELTLA-----QEARAAEKNIUDK 1284

RESULT 7

MYSS\_CYPCA STANDARD; PRT; 1935 AA.

ID MYSS\_CYPCA

AC Q90339;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, fast skeletal muscle.

OS Cyprinus carpio (Common carp).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Cyprinus.

OX NCBI\_Taxid=7962;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fast muscle;

RX MEDLINE=97352533; PubMed=9208928;

RA Hirayama Y., Watabe S.;

RT "Structural differences in the crossbridge head of temperature-

RT associated myosin subfragment-1 isoforms from carp fast skeletal

RT muscle.";

RL Eur. J. Biochem. 246:380-387(1997).

RN [2]

RP SEQUENCE OF 981-1935 FROM N.A.

RC TISSUE=Fast muscle;

RX MEDLINE=97176447; PubMed=9023993;

RA Imai J., Hirayama Y., Kikuchi K., Kakinuma M., Watabe S.;

RT "cDNA cloning of myosin heavy chain isoforms from carp fast skeletal

RT muscle and their gene expression associated with temperature

RT acclimation.";

RL J. Exp. Biol. 200:27-34(1997).

RN [3]

RP SEQUENCE OF 1387-1528 FROM N.A.

RX MEDLINE=95194396; PubMed=787920;

RA Watabe S., Imai J., Nakaya M., Hirayama Y., Okamoto Y., Masaki H.,

RA Uozumi T., Hiroo I., Aoki T.;

RT "Temperature acclimation induces light meromyosin isoforms with

RT different primary structures in carp fast skeletal muscle.";

RL Biochem. Biophys. Res. Commun. 208:118-125(1995).

RL [4]

CC -1- FUNCTION: MUSCLE CONTRACTION.

CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).

CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING

CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,

CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY

CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.

CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT

CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE

CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED

CC SUBFRAGMENT (S2).

CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
DR EMBL: D89992; BAA22069.1; -;  
DR EMBL: D50476; BAA09069.1; -;  
DR EMBL: D43700; BAA07802.1; -;  
DR HSSP: P08799; 1MMD.  
DR InterPro: IPR000048; IO.  
DR InterPro: IPR000009; Myosin\_N.  
DR InterPro: IPR002928; Myosin\_tail.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IO; 1.  
DR Pfam: PF00063; myosin\_head; 1.  
DR Pfam: PF02736; Myosin\_N; 1.  
DR Pfam: PF01576; Myosin\_tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR PRODOM: PD000355; myosin\_head; 1.  
DR SMART: SM00015; IO; 1.  
DR SMART: SM00242; MYSC; 1.  
DR PROSITE: PS50096; IO; 1.  
KW Myosin; muscle protein; coiled coil; thick filament; Actin-binding;  
KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;  
KW Multigene family.  
FT DOMAIN 1 781 MYOSIN HEAD-LIKE.  
FT DOMAIN 782 811 HINGE.  
FT DOMAIN 812 839 COILED COIL (POTENTIAL).  
FT DOMAIN 840 1935 ATP (POTENTIAL).  
FT NP\_BIND 178 185 ACTIN-BINDING.  
FT DOMAIN 659 681 ACTIN-BINDING.  
FT DOMAIN 761 775 METHYLATION (TR1-) (POTENTIAL).  
FT MOD\_RES 129 129 METHYLATION (SH-1).  
FT MOD\_RES 699 699 ALKYLATION (SH-2).  
FT MOD\_RES 709 709 ALKYLATION (SH-2).  
SQ SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;

Query Match 7.3%; Score 173.5; DB 1; Length 1935;  
Best Local Similarity 22.3%; Pred. No. 0.16;  
Matches 72; Conservative 66; Mismatches 126; Indels 59; Gaps 11;

QY 31 LQCLIQSETPAPSRCTPCQRCIQVGKRTITINKLFFDLAODEENVDREF----- 78  
DB 1144 LEETSERLEEGAGTAAQIEENKREAEFCMRDL--EESTLQHEXTAALRKQADS 1200  
QY 79 ---LKNELDNVRAQLSKDKDEKRSOVIDTLRLTLEERNNTVVSLOALCKAEMLCSTL 135  
DB 1201 VAEIGEDIDNQVRKQKLEKESYKMEIDLTSNME-----AVAAKANKLEKCRTL 1253  
QY 136 KQKKYLEQOQDETQKQAEAGRLRSKMT-----MEQIEILLISOLP----- 178  
DB 1254 EDQLSEIKTKSDENVRLQNDMNAQRARLQTFNGEFSRLKEKALV-SQLRKGQATYQQ 1312  
QY 179 -----EVEEMIRDMGQGVSAVEDLAVYCVSLKREYENLKKARRASGEVADKLKDLFSS 232  
DB 1313 IEELKRRIIEEVRKAKNALAHVGSARHDCDLRLROYEEQGRKAKAELRGMSKANSEVAQW 1372  
QY 233 RSKLOT-----VYSLDQAKLELKSQKDLOSADKEIMSLKKKLMLQETLNLPPVASETV 288  
DB 1373 RTKEETAIAIQTLEELKPKKL--AQR-LDAEESIAVNSKASLEKTYQ--RLQGEVE 1427  
QY 289 DRV-LESPAPVEYNLKRPSF 310  
DB 1428 DLMIDVERANSIANDLKKQKQNF 1450

RESULT 8  
MYH4\_RABBIT STANDARD; PRT; 1938 AA.  
AC Q2864; 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin heavy chain, skeletal muscle, juvenile.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NEW ZEALAND WHITE; TISSUE=Skeletal muscle;  
RA Meda K., Hostilnova E., Roesch-Kleinhauf A., Schuster H., Gasperik J.,  
RA Wittlinghofer A.,  
RT "Isolation, sequencing of myosin heavy chain cDNA from rabbit  
RT skeletal muscle and a novel cosynthesis of S-1 fragment with the  
RT essential and regulatory light chains".  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MUSCLE CONTRACTION.  
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES.  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBFRAGMENT (S2).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

FT MOD.RES 708 708 ALKYLATION (SH-2) (BY SIMILARITY).  
SQ SEQUENCE 1938 AA: 223064 MW: D8A8A2EC5B182626 CRC64;  
Query Match 7.2%; Score 172; DB 1; Length 1938;  
Best Local Similarity 21.3%; Pred. No. 0.18;  
Matches 86; Conservative 82; Mismatches 148; Indels 88; Gaps 16;  
OY 31 LOCILISFETAPSRTPGRCIGVGRKRTIKLFPEDLAQEEENVLDRF----- 78  
DB 1146 LEISERLEEGAGTSAQIEMNKKREAFQCKMRDL---EATLQHEATATIRKKHADS 1202  
OY 79 ---LKNELDNVRAQSQDKREKRDQVIITDLRDTLEERNATVVSQQAQAKAEMLCSTL 135  
DB 1203 VAEIGBOIDMLQVQKQLEKEKSELMKEIDDLASNME-----TVSKAKGNLEKKMCRIL 1255  
OY 136 KQOKATLEQOQDETK-----QAQEEAGRL-----RSKAKTMEQI 169  
DB 1256 EDQVSELKTRKEEHQRLINDLSAQARARLQTESEGFSRQDEKDSLVSQLSRGQAFTQOI 1315  
OY 170 ELLQSQLEPEEEMIDMGVQSAVEQLAVYCVSLKKEVENLKEARKASGEVADKLKRL 229  
DB 1316 EEELK-----QLEEEIKAKSLAHALOSARHDCDLRQYEEQEAKELORAMSKANSEV 1371  
OY 230 FSSRSKIQGT---VSELDQAKLELKSQKDLQSAQKKEIMSLKKLTMLOETLNPVAS 285  
DB 1372 AQRTKYETDAIQTELEBAKKL--AQR-LQDAEHEVAVNAKASLEKTKQ--RLQN 1426  
OY 286 EYWDRLV-LESAPAVEVNLKLRPSFRDI-----DLNATFVDVTPPAPSSSQ-- 333  
DB 1427 EVEDLMDIVERTNAACALADKQRNF-DKLIAMKHYETHALELSQKESRSLSTEVP 1485  
OY 334 ---HGVEKL---CLEKSHSPIQ---DVKKIKCKPRKESOL 366  
DB 1486 KVNNAVEESLDQLETLKREKNLQOEISDULEQIABGKRIHEL 1529  
RESULT 9  
MYS2.DICDI  
ID MYS2.DICDI STANDARD; PRT; 2116 AA.  
AC P08799;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Myosin II heavy chain, non muscle.  
GN MHCA.  
OS Dictyostellium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.  
OX NCBI\_TaxId=44689;  
RN RP SEQUENCE FROM N.A.  
RX MEDLINE=87092266; PubMed=3540939;  
RA Warrick H.M., de Lozanne A., Leinwand L.A., Spudich J.A.;  
RT "Conserved protein domains in a myosin heavy chain gene from  
RT Dictyostellium discoideum.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).  
RN RP PHOSPHORYLATION SITES, AND MUTAGENESIS.  
RC STRAIN-AX2;  
RX MEDLINE=90353583; PubMed=2387408;  
RA Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,  
RA Gerisch G.;  
RT "Replacement of threonine residues by serine and alanine in a  
RT phosphorylatable heavy chain fragment of Dictyostellium myosin II.";  
RL FEBS Lett. 269:239-243(1990).  
RN RP PHOSPHORYLATION SITES.  
RX MEDLINE=88112226; PubMed=2828113;  
RA Wagle G., Noegel A., Scheel J., Gerisch G.;  
RT "Phosphorylation of threonine residues on cloned fragments of the  
RT Dictyostellium myosin heavy chain.";  
RL FEBS Lett. 227:71-75(1986).  
RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.  
RX MEDLINE=95345066; PubMed=7619795;  
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,  
RA Rayment I.;  
RT "X-ray structures of the myosin motor domain of Dictyostellium  
RT discoideum complexed with MgADP.Befx and MgADP.AlfA-";  
RL Biochemistry 34:8960-8972(1995).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.  
RX MEDLINE=95345067; PubMed=7619796;  
RA Smith C.A., Rayment I.;  
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the  
RT truncated head of Dictyostellium discoideum myosin to 2.7-A  
RT resolution.";  
RL Biochemistry 34:8973-8981(1995).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
RX MEDLINE=96206189; PubMed=8611530;  
RA Smith C.A., Rayment I.;  
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the  
RT Dictyostellium discoideum myosin motor domain to 1.9-A resolution.";  
RL Biochemistry 35:5404-5417(1996).  
RN [7]  
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.  
RX MEDLINE=97452580; PubMed=9305951;  
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;  
RT "X-ray structures of the MgADP, MgATPamas, and MgAMPNP complexes  
RT of the Dictyostellium discoideum myosin motor domain.";  
RL Biochemistry 36:11619-11628(1997).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.  
RX MEDLINE=98070605; PubMed=9405148;  
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;  
RT "X-ray crystal structure and solution fluorescence characterization  
RT of Mg.2(3')-O-(N-methylanthraniloyl) nucleotides bound to the  
RT Dictyostellium discoideum myosin motor domain.";  
RL J. Mol. Biol. 274:394-407(1997).  
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE  
CC ACTIVITY THAT IS ACTIVATED BY ACTIN.  
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES  
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI  
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS  
CC (MLC-2).  
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL  
CC COREX.  
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN BE FURTHER  
CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
CC SUBFRAGMENT (S2).  
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CHARACTRISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES  
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.  
CC -1- MISCELLANEOUS: DICTYOSTELLUM MYOSIN II HAS NO K(2)EDTA ATPASE  
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1  
CC POSITION (688).  
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 TO DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL: M14628; AAA33227.1; -  
DR PIR: A26655; A26655.  
DR PIR: S00250; S00250.  
DR PDB: 1MMD; 17-AUG-96.

DR PDB: 1MKG: 03-DEC-97.  
 DR PDB: 1MNN: 03-DEC-97.  
 DR PDB: 1MND: 17-AUG-96.  
 DR PDB: 1MNE: 17-AUG-96.  
 DR PDB: 1VOM: 23-DEC-96.  
 DR PDB: 1LVK: 28-JAN-98.  
 DR DictyDb: DD01008: mhca.  
 DR DictyDb: IPR000048: IQ.  
 DR InterPro: IPR004009: Myosin\_N.  
 DR InterPro: IPR001609: myosin\_head.  
 DR Pfam: PF00612: IQ; 2.  
 DR Pfam: PF00063: myosin\_head; 1.  
 DR Pfam: PF02736: Myosin\_N; 1.  
 DR PRINTS: PR00193: MYOSINHEAVY.  
 DR ProDom: PD000355: myosin\_head; 1.  
 DR SMART: SM00015: IQ; 1.  
 DR SMART: SM00242: MYSC; 1.  
 DR PROSITE: PS50096: IQ; 1.  
 KM Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;  
 KM Calmodulin-binding; Methylation; Alkylation; Phosphorylation.  
 FT DOMAIN 1 761 MYOSIN HEAD-LIKE.  
 FT DOMAIN 762 791 IQ.  
 FT DOMAIN 817 2116 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP.  
 FT DOMAIN 638 660 ACTIN-BINDING.  
 FT DOMAIN 738 752 ACTIN-BINDING.  
 FT MOD\_RES 130 130 METHYLATION (DI-) (POTENTIAL).  
 FT MOD\_RES 678 678 ALKYLATION (SH-1).  
 FT MOD\_RES 1823 1823 PHOSPHORYLATION (BY MHCK).  
 FT MOD\_RES 1833 1833 PHOSPHORYLATION (BY MHCK).  
 FT MOD\_RES 2029 2029 PHOSPHORYLATION (BY MHCK).  
 SQ SEQUENCE 2116 AA: 243871 MW: 2FC3770B1E56A1 CRC64:

Query Match 7.28; Score 172; DB 1; Length 2116;  
 Best Local Similarity 27.4%; Pred. No. 0.2;  
 Matches 90; Conservative 49; Mismatches 109; Indels 80; Gaps 17;

QY 60 NKLFPDLAQEEENVLD-RELKNELDNVRALQSKD---KEKRDQVITIDTNRTEERN 115  
 DB 852 DKLEKSLKQTESNVDLQRLQKAEKELKAMYSKDALEKQRELETRVEDMESELEDEKK 911  
 QY 116 ATVVSLOALGAKEMLCSTLKQMKYLEQOODETKQAQAEGRLSKMKTMEOITELLLOS 175  
 DB 912 LALENLQNKRSYE-----EKVRDELEELQEE-----QKLRNTELEKAKK 951  
 QY 176 QLPEVEEMIMDVGOS-AVEOLAVYCVSLKKEYENLKEARKASGEVADK--LRKDLFSS 232  
 DB 952 YEELTEEMKR-VNDGQSDTISRLEKIKDELQKEVEELTES--FSEESKDKGVLEK---T 1004  
 QY 233 RSKLQVYVSLDQAKLELSAOKDLSADKEIMSLKKKKLMLOETLNLPPVASETVDRLV 292  
 DB 1005 RVRLO--SELDLTVNLDETQKSELRLQKKLEELQVODAL-----AAETAAKLA 1056  
 QY 293 LESPAPAEVNLKLRSEFRDIDLNATFDVTPPARSSQHGYYEKLCEKSHSPIQDV 352  
 DB 1057 QEA-----ANKKIQ-----GEYTELEKFNSEV--TARSN-----VEKSKTLL-- 1092  
 QY 353 PKKICKPKRESOLSLGQSCAGEPDEE 380  
 DB 1093 -----ESQL-----VAVNNELDEE 1106

RESULT 10  
 ID RA50\_AQUAE STANDARD: PRT: 978 AA.  
 AC 067124:  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Probable DNA double-strand break repair rad50 ATPase.  
 GN RAD50 OR AO\_1006.  
 OS Aquifex aeolicus.

OC Bacteria: Aquificales; Aquificaceae; Aquifex.  
 OX NCBI\_TaxID=63363;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=VF5;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 aeolicus.";  
 RL Nature 392:353-358(1998).  
 CC -I- FUNCTION: Involved in DNA double-strand break repair (DSBR). The  
 CC rad50/mre11 complex possesses single-strand endonuclease activity  
 CC and ATP-dependent double-strand-specific exonuclease activity.  
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding  
 CC and/or repositioning DNA ends into the mre11 active site (By  
 CC similarity).  
 CC -I- SUBUNIT: Forms a complex with mre11 (By similarity).  
 CC -I- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE000718; AAC07092.1; -  
 DR InterPro: IPR003439: ABC\_transport.  
 KM DNA repair: Hydrolyase; ATP-binding; Coiled coil; Complete proteome.  
 FT NP\_BIND 32 39 ATP (BY SIMILARITY).  
 FT DOMAIN 160 826 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 978 AA: 115897 MW: 9B0F2BF51ADD1151 CRC64:

Query Match 7.1%; Score 170; DB 1; Length 978;  
 Best Local Similarity 22.6%; Pred. No. 0.11;  
 Matches 79; Conservative 70; Mismatches 125; Indels 76; Gaps 13;

QY 41 APSKTCPOCRITQVGRKTIIN-----KLFPDLAQEEENVLDREF-----LKNE 82  
 DB 481 SPGDYCPGCGIRGKALEVNDAGISELKHAKLEKKEKEIDTTLKVAQKINSKKEE 540  
 QY 83 LDNVRALQSKDKKRDQVYITLDRDTEERNATVVSLOALGAKEMLCSTLKQMKYL 142  
 DB 541 MEKLIRNEVEELRKE-----IPENLKER--IKLEELRIEKEKLEHKLNKYRKAL 587  
 QY 143 EQOODETKQAQAEGRLSKMKTMEOITELLQSOI--PEVEEMIRDMGVGSAVEOLAVYC 201  
 DB 588 EDKQKQEEQAKLHKRQTELELKE-KIREKSLVKEFEELR-----VERLEDYE 638  
 QY 202 VSLKKEYENLKEARKASGEVADLRK--DLFSSRSKLQVYSELDOAKLELSAOKDLO 258  
 DB 639 ESLKEELINYSNLQLELEKEKRLKRFEELSRKSLSLE-----ELALNESIN 688  
 QY 259 SADKEIMSLKKTLMLOETLNLPPVASETVDRLVLESAPVE-----VNLKLRP 308  
 DB 689 SLEEKREKLELANIYEVAKSP---REVVELYLGQKEALEKRIKFEESFOSLKLKS 745  
 QY 309 SFRDDIDLNATFDVTPPARSSQHGYYE--KLCEKSHSPIQDVPKTI 356  
 DB 746 EIEEKLK-----EYEGIRELSDINGEYSVKTOLEKHKLGAEVKREL 788

RESULT 11  
 ID TANA\_XENLA STANDARD: PRT: 1744 AA.  
 AC 001550:  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)







[illegible]

```

FT FT DOMAIN <1 619 GLOBULAR HEAD.
FT DOMAIN 620 912 COILED COIL (POTENTIAL).
FT DOMAIN 917 941 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 922 922 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 941 AA; 105152 MW; AB5C10895CD7EE08 CRC64;

Query Match 7.0%; Score 168; DB 1; Length 941;
Best Local Similarity 19.6%; Pred. No. 0.13;
Matches 74; Conservative 55; Mismatches 96; Indels 152; Gaps 10;

QY 52 QVGRKTIINKLFF-----DLAQEEENVLDREFLK--NELDNV--RAQDS 91
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 553 RIGKENFLEKLGFTISKHNLYSRASQKRPDPNPSPRYMFDHEFTLVKLEGVITKATYK 612
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 92 QDKDKRDSOVIIDTLRPTLEERNATVYSLOALSKAMLSSTLKKOKKYLEEQODE-- 148
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 613 SSEEKKEEYV-----KTLLEOHNDIVITHYKMIITEQDLQLELKQOYSTLKCOMBOLOT 667
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 -TKQAQ-----EEAGRLRSKMTMBQI 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 668 AVTQQAQSIQGHKDDYNLLKVLQDGKDNHOGSHDGAQVNGIQREISRLREIRELSQ 727
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 ELLLOSQREPEVEMTRDMGVGQSA--VQDLAVYC----- 201
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 QALLQGLAKEDSLIENLKSSQASGMSQASQATCPPRDPQVAYELKQELTALKSOLCSQS 787
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 202 -----VSLKKEYENLKARK 216
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 LEITTLQTEENCELLQRAETLAKSVPEGESEHVSAAKTTDVEGRLSALLQETKELMEIK 847
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 217 ASGEVADLRKLDFFSSRSKLTQTVSELDQAKLELKSAAKD-----LOSADKEITMSLK 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 848 ALSEERTAIQKQKDDSSNSTIATLQTEKPKLDLEVIYDSKKEQDDLVLADQDKILSLKS 907
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 270 KITMLQETLMLPVPASE 286
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 908 KKLKDLGH-----PVVEE 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
G160_MOUSE STANDARD: PRT: 1325 AA.
AC P55937:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Golgin-160 (Male-enhanced antigen-2) (MEA-2).
GN GOLGA3 OR MEA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Testis;
RX MEDLINE=97217683; PubMed=9063644;
RA Kondo M., Sutoh S.,
RT "Cloning and molecular characterization of cDNA encoding a mouse
RT male-enhanced antigen-2 (Mea-2): a putative family of the Golgi
RT autantigen."
RL DNA Seq. 7:71-82(1997).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY
CC DETECTABLE MALE ANTIGEN (SDM).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN
CC LEYDIG CELLS, SPERMATOZOONIA, OR SPERMATOCYTES.
CC -1- SIMILARITY: HIGH, TO HUMAN GOLGIN-160.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE
CC INITIATOR.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

```



```

Db 1149 LEEISERLIEBAGGATSAQIEMNKKREAEFOKMRDL---EEATLQHEATATLKRKHADS 1205
QY 79 ---LKNELDNVRAQLOSKOKKEKRDISOYIITDLETLERNATVVSLOALGKEMLCSTL 135
Db 1206 VAEIGQIDMLQHVQKOKLEKSEKSEMDIDLASNVE-----TVSKAGNLEKMCRTL 1258
QY 136 KKGMYLLEQOQDETK-----QAQEEAGRL-----RSKMKTMEOI 169
Db 1259 EDOLSELKSEEEBOQRILNDLTQGRGLQTESESEFSDOLDEKALVSQLSRGQAFTQOI 1318
QY 170 ELLLOQLPVEVEEMIDMGVGSVAEQLAVYCSLKEKENLKEARKASGEVADKLKLDL 229
Db 1319 EEELKR-----QLEEEIKAKNALAHALQSSRHDCDLREQYEEQESKAELOALSKANTEV 1374
QY 230 FSSRSKQOT-----VYSLOQAKLELSAQKDLSADKEINSLKKKLTMLDET 277
Db 1375 AQWRTYETDAIQRTLELEAKKRL--AQR-LQAEEHVAEAVNAKCASLERT 1423

```

```

RESULT 17
MISP_ONCVO STANDARD; PRT; 879 AA.
ID MISP_ONCVO
AC 002171;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Paramyosin.
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93165084; PubMed=7679477;
RA Dahmen A., Gallin M., Schumacher M., Ertmann K.D.;
RT "Molecular cloning and pre-mRNA maturation of Onchocerca volvulus
  paramyosin.";
RL Mol. Biochem. Parasitol. 57:335-338(1993).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
  MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M95813; AAA29431.1; -.
CC HSSP: P80220; IDIP.
CC InterPro: IPR002928; Myosin_tail.
CC Pfam: PF01576; Myosin_tail; 1.
CC Coiled coil, muscle protein; Thick filament; Myosin.
CC FT DOMAIN 1 36 NONHEMICAL REGION (POTENTIAL).
CC FT DOMAIN 37 860 COILED COIL (POTENTIAL).
CC FT DOMAIN 861 879 NONHEMICAL REGION (POTENTIAL).
CC SQ SEQUENCE 879 AA; 101125 MW; DCA24AC01A9F02B1 CRC64;

```

```

Query Match 7.08; Score 167; DB 1; Length 879;
Best Local Similarity 19.88; Pred. No. 0.13;
Matches 105; Conservative 88; Mismatches 160; Indels 178; Gaps 20;

```

```

QY 59 INKLFPLDQ-----EEVVLDERF--LKNELDN--VRAQLOSKOKKRR----- 98
Db 198 LNRHVNDLDAQQRRLQALQENNDLKEITHQVQDNLQHVAYQLAQDLLEARRRLEDAERE 257
QY 99 -----DSQVLIIDTLRLDLEERNATVVSLOALGKA----- 128

```

```

Db 258 RSQLOAQLOHVQLELDVSFTALDEESARAARAEHKLALANTETITQKSKDAEVALHHEE 317
QY 129 --EMLCSTLKKQKMYLEQOO---DETKQAOEEBGRLSKMKTY----- 166
Db 318 VEDLRKMKLOKQAEVEIEMLOKLSOLEKAKSRLOSEVEVLIDLEKNAONTAILERA 377
QY 167 -EOIELLOQLPVEVEEMIDMGVGSVAEQLAVYCSLKEKENLKEARKASGEVADKLKLDL 207
Db 378 KEQLEKTVNELKTRIDELVELYEAQREARALAELOKKKNLYEKAVEQKALARENKYL 437
QY 208 YENLKEARKASGEVADKLKRL-DEFSR-----SKLOTYVSEI 243
Db 438 QDOLHEAKKEALDANKRLHELDLENRLAGEIRELOTLAKESPAARRDAENRQALAEI 497
QY 244 DQAKLELSAQKDLSADKEINSLKKKLTMLDETLPVPASTVRL--VLESAPAE 300
Db 498 QOLRIEM--ERRLOKEEEMELRKNMOP-----EIDRLTAALDAARAKK 541
QY 301 VNLKLRSPSPFDI-DLNATFD-----VDTPPARSSOHGYEKEKLEKSHSPIQDVP 353
Db 542 AEIARLKKKYQAEIAELEMTVDNLNANIFAOKTIRKQSOQLKVLQASLEDTORQLOTL 601
QY 354 KKICKGRKESQSLGQSCAGEPDELVGAFPIFYRNAILGQKPKRPRESSCSKDVV 413
Db 602 DQYALQQRKVSALSALAELEC-----KVALDNALIRAKQAEIDLEAN----- 643
QY 414 RTGFDGLGGRKTFQIPDVTYMIPLPYKPKTKQKQKRVKTVSLSQAKLD 464
Db 644 -----GRI-----TDIVSIN-----NNLTAKN--KLETSLSTAQADLD 675

```

```

RESULT 18
MISP_BRUMA STANDARD; PRT; 880 AA.
ID MISP_BRUMA
AC 001202; P90711;
DT 01-FEB-1996 (Rel. 33, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Paramyosin.
OS Brugia malayi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N.A.
RA Langy S., Luguid P., Nicolas L.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RX MEDLINE=92131068; PubMed=175173;
RA Li B., Chandrasekhar R., Alvarez R.M., Lifits F., Weil G.J.;
RT "Identification of paramyosin as a potential protective antigen
  against Brugia malayi infection in jirds.";
RL Mol. Biochem. Parasitol. 49:315-324(1991).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
  MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U77590; AAC18613.1; -.
CC EMBL: M63097; AAA27859.1; -.
CC EMBL: M63098; AAA27860.1; -.
CC HSSP: P80220; IDIP.

```

```

DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail: 1.
KM Coiled coil; Muscle protein; Thick filament;
FT DOMAIN 1 34 NONHELICAL REGION (POTENTIAL).
FT DOMAIN 35 859 COILED COIL (POTENTIAL).
FT DOMAIN 860 880 NONHELICAL REGION (POTENTIAL).
FT CONFLICT 189 189 V -> I (IN REF. 2).
FT CONFLICT 197 197 L -> A (IN REF. 2).
FT CONFLICT 251 251 L -> F (IN REF. 2).
FT CONFLICT 255 255 E -> Q (IN REF. 2).
FT CONFLICT 696 696 A -> R (IN REF. 2).
FT CONFLICT 736 736 FRKLRR -> IQEATPA (IN REF. 2).
FT CONFLICT 825 825 L -> IQ (IN REF. 2).
FT CONFLICT 857 880 HOLLRKMLOROKETFEKMSNRDN -> SSVVTKGNASAK
FT IYVEDRO (IN REF. 2).
SQ SEQUENCE 880 AA; 101904 MW; F79A27E642D4A85E CRC64;

Query Match 7.0%; Score 167; DB 1; Length 880;
Best Local Similarity 19.6%; Pred. No. 0.13;
Matches 104; Conservative 89; Mismatches 160; Indels 178; Gaps 20;

QY 59 INKLFEDLADE-----EENVLDREF--LKNELDN---VRAQLSQDKDKR----- 98
D 198 LNRHVNDLAQORQLQAENNNDLKEIHQKQVLDNLQHVQYLAQOLEEARRLEDAERE 257
QY 99 -----DSQVIIDPLRDTLEERNATVVSLSQALGKA----- 128
D 258 RSQQAQOLHOVQVLELDVFRFALDEESAPARAEEKRLALANTETIQWMSKFDPAEVALHHEE 317
QY 129 --EMILCSTLKQKMYLBOQ---DETROADEAGRLSKMKT----- 166
D 318 VEDRRKMKLOQAQAYEQIEIMLOKISQLEKAKSRLOSEVVLVDEKAQNTAILERA 377
QY 167 -EQLELLQSLPVEVEEMIRMGVQ-----SAVEQLAVYCVSLKKE 207
D 378 KEQLKTVNELKVIIDELTYELKAEQARAALAELOKLNLYEKAVEQKALARENKKL 437
QY 208 YENLEKARKSGEVAADKLK--DLFSSR-----SKLTQVYSEL 243
D 438 ODDHEKEKALADANKRLHELDLENARLAGIRELOTLAKSEARDAENKRAQALAEEL 497
QY 244 DQAKLEKSAQKDLQSADEKELMSLKKLTMLQETLNPVAVSETVDRL--VLESPARVE 300
D 498 QQLRTM---ERRLOKEEKEEMALRKMQF-----EIDRLTALADAEARMK 541
QY 301 VNKLKRRPFRDDI--DLNATFD-----VDPPARPSSQHGVEKLCLEKSHSPIQDVP 353
D 542 AEISRLKRRKYQAEIAELMTVDNINRANIEAQTKIKQSEQLKILQASLEDTOROLOOTL 601
QY 354 KIKCGPKRESQSLGSCSGEPDEELVGAFPIFVRNALIGOKPRPRESSESCSDVY 413
D 602 DQVLAQRKVSALSAELEEC-----KVALDNALFARKQAEIIDELEAN----- 643
QY 414 RTGEGDLGGRTKTIQPTDYVIRPLPVKPKTKVQRVARKVTPSLFQAKID 464
D 644 -----GRF-----TDLVSVN-----NNLTAKN--KLETELSTAQADLD 675

RESULT 19
MYSA_DROME STANDARD; PRT; 2411 AA.
AC P05661;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, muscle.
GN MHC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```

```

OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89384556; PubMed=2506434;
RA George E.L., Ober M.B., Emerson C.P. Jr.;
RT "Functional domains of the Drosophila melanogaster muscle myosin
RL heavy-chain gene are encoded by alternatively spliced exons."
RN Mol. Cell. Biol. 9:2957-2974(1989).
RP [2]
RX SEQUENCE OF 1-312 FROM N.A.
RP MEDLINE=87280141; PubMed=3038896;
RA Wassenberg D.R. II, Kronert W.A., O'Donnell P.T., Bernstein S.I.;
RT "Analysis of the 5' end of the Drosophila muscle myosin heavy chain
RT gene. Alternatively spliced transcripts initiate at a single site and
RT initiate locations are conserved compared to myosin genes of other
RT organisms."
RL J. Biol. Chem. 262:10741-10747(1987).
RN [3]
RP SEQUENCE OF 486-881 FROM N.A.
RP STRAIN=CANTON-S; TISSUE=Embryonic muscle;
RX MEDLINE=91330870; PubMed=1907912;
RA Kronert W.A., Edwards K.A., Roche E.S., Wells L., Bernstein S.I.;
RT "Muscle-specific accumulation of Drosophila myosin heavy chains: a
RT splicing mutation in an alternative exon results in an isoform
RT substitution."
RL EMBO J. 10:2479-2488(1991).
CC - FUNCTION: MUSCLE CONTRACTION.
CC - SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC - SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC - ALTERNATIVE PRODUCTS: AS DROSOPHILA HAS A SINGLE MUSCLE MHC GENE,
CC MHC ISOFORMS MUST RELY ON ALTERNATIVE RNA SPLICING THAT LEADS TO
CC DIFFERENCES IN THE C-TERMINUS OF THE VARIOUS MHC PROTEINS.
CC - MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC - SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC - SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M61229; AAA28686.1; ALT_SEQ.
DR EMBL; M61229; AAA28687.1; ALT_SEQ.
DR EMBL; J02788; AAA28706.1; ALT_SEQ.
DR EMBL; J02788; AAA28707.1; ALT_SEQ.
DR EMBL; X60196; CAA42752.1; ALT_SEQ.
DR EMBL; X60196; CAA42753.1; ALT_SEQ.
DR EMBL; X60196; CAA42754.1; ALT_SEQ.
DR PIR; A28492; A28492.
DR PIR; A32491; A32491.
DR PIR; B32491; B32491.
DR HSP; P08799; LMND.
DR FlyBase; FBgn0002741; Mhc.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR Pfam; PF00063; myosin_head; 7.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 5.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.

```

DR PROSITE; PSS0096; IQ; 1.  
 KM Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KM ATP-binding; Calmodulin-binding; Alternative splicing;  
 KM Multigene family.  
 FT DOMAIN 1 2  
 FT DOMAIN 1205 1232 MYOSIN HEAD-LIKE.  
 FT DOMAIN 1233 1236 IQ.  
 FT DOMAIN 1237 1237 COILED COIL (POTENTIAL).  
 FT N-BIND 227 234 ATP (BY SIMILARITY).  
 FT VARSPIC 2385 2385 P -> I (IN SHORT ISOFORM).  
 FT VARSPIC 2386 2411 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 43 44 EK -> RE (IN REF. 2).  
 SQ SEQUENCE 2411 AA; 276434 MW; 8C4FEC1BB8D310A9 CRC64;

Query Match 7.0%; Score 167; DB 1; Length 2411;  
 Best Local Similarity 20.5%; Pred. No. 0.42; Mismatches 144; Indels 140; Gaps 15;  
 Matches 93; Conservative 76;

QY 66 LAQEEENVLDREFLNELNDVRAQLSQKDEK-----RDSQVITDITRDITL----- 112  
 DB 1999 LEQENNVLAQL---ELSGVROEIDRIQEKKEEFENTKKNHQRALDSQASLEAKG 2055  
 QY 113 -----ERNATVVSLSQALGKAEMLCSTLKKQKMYLEQODETKQAQDEAGRLSKM 163  
 DB 2056 KAEALRRKKLEADINLELALDHANKANAQAQKIRYQOOLKIDIOALTEEEQRARDDA 2115  
 QY 164 KITMOIQL-----LQSOLPEVEEMIRDMGVGSAVEQLAVYVSLKKEVNEKARKA 217  
 DB 2116 R-EQLISERRANALONEESRTLEQADRGRAQ-----ELADAEQJNEVSAQ 2167  
 QY 218 SGEVADLRKDLFSSRSKLTQTVSELDQ-----AKL--ELKSAQKD 256  
 DB 2168 NASSIAKRR---LESELQTHSDDELINLEANKSEKAKKANVADARLADDELRAQDH 2223  
 QY 257 LQSDAKTMSLKKKLYLQETLNPV-----ASEYDRLVLE-----SPAP 298  
 DB 2224 AQTOEKLRKALEQIKELQYRLDEAEANALKGKKAIOKLEQRYVELNELDGEQRHAD 2283  
 QY 299 VEVL-----KLRRSPRDOIDLNAFTDVPDPAPRSSQHGVEKLEKLSHSPIQDVP 353  
 DB 2284 AQKNLRKSERVKELSFQSEED-----RKNHRRMODLV 2316  
 QY 354 KIKCKGPR-----KESQLSGQSCAGPEDELVGAFPIFVENAI 393  
 DB 2317 DKLOQKIKTKYKQTEAEAEIALNLAKFRKAQQLFEAEERADLAQALISFRKAGNAGS 2376  
 QY 394 LGQKQPKRPSSESCSDVYRTGDTGLGGRKF 426  
 DB 2377 VGRGASPAFRATS-----VRPQFDGLAFPPRF 2403

RESULT 20  
 PAM\_STRPY STANDARD; PRT; 388 AA.  
 AC P49054;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Plasmidogen-binding protein pam precursor (Fragment).  
 GN PAM.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxID=1314;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 32-41.  
 RC STRAIN=AP53 / Serotype M53;  
 RX MEDLINE=94064605; Pubmed=8244975;  
 RA Berge A., Sjoerding U.;  
 RT "PAM, a novel plasmidogen-binding protein from Streptococcus pyogenes."  
 RU J. Biol. Chem. 268:25417-25424(1993).  
 CC -I- FUNCTION: BINDS TO HUMAN PLASMINOGEN. COULD PROVIDE THE BACTERIA

CC WITH A MECHANISM FOR INVASION.  
 CC -I- SUBCELLULAR LOCATION: CELL WALL ASSOCIATED.  
 CC -I- SIMILARITY: TO OTHER STREPTOCOCCAL AND STAPHYLOCOCCAL PROTEINS IN THE REGION OF THE MEMBRANE ANCHOR.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: Z22219; CAAB0222.1; -  
 DR HSSP: P03069; IGCL.  
 DR InterPro: IPR001899; Gram\_pos\_anchor.  
 DR InterPro: IPR003345; M\_repeat.  
 DR Pfam: PF02370; M; 8.  
 DR PROSITE: PSS00343; GRAM\_POS\_ANCHORING; 1.  
 KM Signal: Repeat; Cell wall.

FT SIGNAL <1 19 OR 28 (POTENTIAL).  
 FT CHAIN 20 >388 PLASMINOGEN-BINDING PROTEIN PAM.  
 FT DOMAIN 91 116 2 X APPROXIMATE TANDEM REPEATS, TYPE A.  
 FT REPEAT 91 103 A-1.  
 FT REPEAT 104 116 A-2.  
 FT DOMAIN 147 161 2 X TANDEM REPEATS, TYPE B.  
 FT REPEAT 147 153 B-1.  
 FT REPEAT 154 161 B-2.  
 FT REPEAT 163 278 3 X TANDEM REPEATS, TYPE C.  
 FT REPEAT 163 204 C-1.  
 FT REPEAT 205 246 C-2.  
 FT REPEAT 247 278 C-3 (INCOMPLETE).  
 FT REPEAT 344 380 GLX/PRO-RICH.  
 FT DOMAIN 381 386 CONSERVED IN GRAM-POSITIVE COCCI SURFACE  
 FT PROTEINS.  
 SQ NON\_TER 388 388  
 SQ SEQUENCE 388 AA; 43629 MW; EEECAF9D62CCDB12 CRC64;

Query Match 7.0%; Score 166; DB 1; Length 388;  
 Best Local Similarity 20.2%; Pred. No. 0.059;  
 Matches 73; Conservative 62; Mismatches 120; Indels 106; Gaps 11;

QY 65 DLQEEENV-----LDREF--LKN-----ELDVNRALQSKQKDEKRSQVITD 105  
 DB 70 DLREKEELQCKDVERKLTADAELORLKNERHEAELEKRSRHDHDKAEARKALED 129  
 QY 106 TLRTDLEERNATVVSLSQALGKAEMLCSTLKKQKMYLEQODETKQAQDEAGRLSKMT 165  
 DB 130 KLADKQEHNGA-----LRYINEKEAEKEKEAE---QKTKLE 164  
 QY 166 MEQIELLSQSLPEVEEMIRDMGVGSAVQLAVYCVSLKKEVNEKAEKASGEVADKL 225  
 DB 165 EKQISDASROGLR-----BDLDASREARKQVEKDLNLTAEIDKVKAEKQISDASROGL 218  
 QY 226 RKDLFSSRSK-----LQTVYSELDAQLE-----LKSQAKDL 257  
 DB 219 RRDIDASREARKQVEKGLANTAEIDKVKAEKQISDASROGLRDLASREARKQVEKAL 278  
 QY 258 QSADEKIMSLSKKLTMLQETLNPVASETVDRLVLESPAPVEVNLKLRSPFRDDIDLN 317  
 DB 279 EANSKTLAEKLNKLELES-----KLTKEKKEKELQAKLEAEKALKLEQALAK 327  
 QY 318 A-----TFDVTTPPARPSSQHGVEKLEKLSHSPTIDVPPKICKGPKESQLS 367  
 DB 328 AEELAKLRAEKASDQTPDAPKPN-----KAVPGGGAQOAGCTKPNQNKAPKMETKRQ 380  
 QY 368 L 368  
 DB 381 L 381

RESULT	21			
ID	VDP_RAT	STANDARD:	PRT:	959 AA.
AC	P41542;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	General vesicular transport factor p115 (Transcytosis associated protein) (TAP) (Vesicle docking protein).			
GN	VDP.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eultheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxId=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=95132632; PubMed=7831123;			
RA	Barroso M., Nelson D.S., Szul E.;			
RT	"Transcytosis-associated protein (TAP)/p115 is a general fusion factor. Required for binding of vesicles to acceptor membranes."			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:527-531(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	TISSUE=Liver;			
RX	MEDLINE=95132633; PubMed=7831124;			
RA	Barroso M., Nelson D.S., Szul E.;			
RT	"Transcytosis-associated protein (TAP)/p115 is a general fusion factor. Required for binding of vesicles to acceptor membranes."			
RL	Proc. Natl. Acad. Sci. U.S.A. 92:527-531(1995).			
CC	-I- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR INTERCELLULAR TRANSPORT IN THE GOLGI STACK: IT IS REQUIRED FOR TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR AND TARGET MEMBRANES IN PROXIMITY.			
CC	-I- SUBUNIT: HOMODIMER WITH TWO N-TERMINAL HEADS AND A C-TERMINAL COILED-COIL TAIL. DIMER FORMED BY PARALLEL ASSOCIATION OF THE TAILS.			
CC	-I- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.			
CC	-I- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.			
CC	-I- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER; PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS. DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE; PHOSPHORYLATED PROMOTES DISSOCIATION (BY SIMILARITY).			
CC	-I- SIMILARITY: BELONGS TO THE VDP/USOL/RLB047c FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; U14192; AAA62632.1; -;			
DR	EMBL; U15589; AAC52151.1; -;			
DR	InterPro: IPR000225; Atmadillo.			
DR	PROSITE: PS50176; ANM_REPEAT; UNKNOWN_1.			
KW	Transport; Protein transport; Golgi stack; Membrane; Coiled coil;			
KW	Phosphorylation.			
FT	DOMAIN	1	637	GLOBULAR HEAD.
FT	DOMAIN	638	930	COILED COIL (POTENTIAL).
FT	DOMAIN	935	959	ASP/GLU-RICH (ACIDIC).
FT	MOD_RSS	940	940	PHOSPHORYLATION (BY SIMILARITY).
FT	CONFLICT	591	591	S -> P (IN REF. 2).
FT	CONFLICT	658	658	M -> V (IN REF. 2).
FT	CONFLICT	816	816	S -> R (IN REF. 2).
FT	CONFLICT	873	873	A -> S (IN REF. 2).
SO	SEQUENCE	959 AA;	107162 MW;	3565994B48C7E003B CRC64;

Query Match	Similarity	20.2%	Score 166	DB 1	Length 959
Best Local	Similarity	20.2% <td>Pred. No. 0.16</td> <td></td> <td></td>	Pred. No. 0.16		
Matches	76	Conservative	56	Mismatches	93
				Indels	152
				Gaps	11
QY	52	QVGRKTIINLKEF-----	DLAQEEENVLDREFLK--NELDNV--RAQLS	91	
DB	571	RIGKENYIEKLFQISNHELVSASQKPPQNPSPSEMFIDHEFTKLVKELGVITKAIVK		630	
QY	92	QNDKEKRSQVYIIDRLDRIEERNATPVVSLQALGAEMILCSTIKKQMYLRLQDQD---	148		
DB	631	SEEDKEKEEV-----KTLTEQHDNIIVTHYKNNIRQDLQLEBKQVSTLKCQNSQLQT	685		
QY	149	--TKQAO-----	BEAGRLSRSMKMTQEI	169	
DB	686	AVTQQAASQIQHKDDQYNLLKVLQKGDNNHQSHSDQAQVNGIQPREISLRBEIELRSH	745		
QY	170	ELLLOSQLPVEEMIRDM-----GVGQSAV-----EQLAIV-----	200		
DB	746	OVLLQSOLAKEQPTVIEHLRSSQVSGMSEQALATCSPRDAEYVAELKQELSAKLSQLCSQS	805		
QY	201	-----	CVSLKEVENLKEARK	216	
DB	806	LEITLQIENSLQQAQRETLAKSVYEGSESLVTAKTIDVGRSLALLQETRELKELNK	865		
QY	217	ASGEVADKLKRDLFSSRSKILQTVSELDQAKLELSAQKD-----LQSADEKIMSLRK	269		
DB	866	ALSEERTAIQKOLDSSNSTIALLQTEKDKLYLEVDSKKEQODLLVLLADQOKILSLKS	925		
QY	270	KLTMLQETLNLPRVASE	286		
DB	926	KLKLDGH-----PVEDE	937		
RESULT 22					
MYH7_PIG STANDARD: PRT: 1935 AA.					
AC	15-DEC-1998	(Rel. 37, Created)			
DT	15-DEC-1998	(Rel. 37, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Myosin heavy chain, cardiac muscle beta isoform (MYHC-beta).				
GN	MYH7.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_taxid:9823;				
RA	SEQUENCE FROM N.A.				
RC	STRAIN-DOMESTICA;				
RL	KO Y.L.;				
CC	Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.				
CC	-1- FUNCTION: MUSCLE CONTRACTION.				
CC	-1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2				
CC	HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)				
CC	AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).				
CC	-1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.				
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING				
CC	CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,				
CC	CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.				
CC	-1- PM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY				
CC	ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.				
CC	-1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT				
CC	MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE				
CC	SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED				
CC	SUBFRAGMENT (S2).				
CC	-1- MISCELLANEOUS: THE CARDIAC ALPHA ISOFORM IS A 'FAST' ATPASE				
CC	MYOSIN, WHILE THE BETA ISOFORM IS A 'SLOW' ATPASE.				
CC	-1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.				
CC	-1- SIMILARITY: CONTAINS 1 IQ DOMAIN.				
CC	-----				
CC	This SWISS-PROT entry is copyright It is produced through a collaboration				



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U75316; AAB37320.1; -  
 DR HSSP: P08799; ILVK.  
 DR InterPro: IPR000048; IO.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IO; 2.  
 DR Pfam: PF00663; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IO; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PSS0096; IO; 1.  
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Multigene family;  
 KW Calmodulin-binding.  
 FT DOMAIN 1 780 MYOSIN HEAD-LIKE.  
 FT DOMAIN 839 1935 COILED COIL (POTENTIAL).  
 FT NP\_BIND 178 185 ATP (POTENTIAL).  
 FT DOMAIN 655 677 ACTIN-BINDING.  
 FT DOMAIN 757 771 ACTIN-BINDING.  
 FT MOD\_RES 129 129 METHYLATION (TRI-) (POTENTIAL).  
 FT MOD\_RES 695 695 ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 705 705 ALKYLATION (SH-2) (POTENTIAL).  
 SQ SEQUENCE 1935 AA; 223110 MW; 149CDBFD910DBB08 CRC64;

Query Match 7.0%; Score 166; DB 1; Length 1935;  
 Best Local Similarity 22.2%; Pred. No. 0.37;  
 Matches 74; Conservative 62; Mismatches 131; Indels 66; Gaps 9;

QY 79 LKNEIDNVAQALQSKDEKRDSDYITLRLDTEE---RNATVVSLOQALGAKMELCST 134  
 DB 1287 LSRQLEDEALISQTRKGLTYTOOLEDLKRLQLEEVYAKNALHALQSAHNAHDLLEQ 1346  
 QY 135 LKQKMTYLEQOODETKQAEAGRLRSKMT-----MQETILLQSQPREVEEMIRD 186  
 DB 1347 YEEETKTAELQRLVLSKANSEVAQWRKYETDAIQRTTELEAKKLAQRLQDAEE--- 1402  
 QY 187 MGVCQSAVEQLAVYCVSLK-----EYENIK---EARKASGEVADKLRLDLFSSRSKL 236  
 DB 1403 -----AVEAVNAKCSSELEKTKHRLQNEIDELMDVENSNAALALDKKQKQNFKILAEW 1456  
 QY 237 QTVVSELDQALTEKSAQDLQADKEIMSLKKL-----TMLQETLNLPPVASETV 288  
 DB 1457 KQKTEE---SQSELESSQKEARSLSTELFKLNAYEESLEHLETSKRNKNMLQEEISDLT 1513  
 QY 289 DRVLESAPPEVNIKLKRPSPFRDDIDNATFDDVTPPARSSQSHGYEKLCEKSSP 348  
 DB 1514 EQLSSSGKTHLEL-KYRKQLEAELELQ-----ALEESES 1550  
 QY 349 IQDVPRKIKCGPRKESQSLGQSCAGPDEEL 381  
 DB 1551 LEHEBGKILRAQLEFNQIKAMEKRLAEKDEEM 1583

RESULT 23  
 MYH1\_HUMAN STANDARD: PRT; 1939 AA.  
 AC P12882; O9Y622;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Myosin heavy chain, skeletal muscle, adult 1 (Myosin heavy chain  
 DE IIX/d) (MyHC-IIX/d).  
 GN MYH1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Skeletal muscle;  
 RX MEDLINE=99318869; PubMed=10388558;  
 RA Weiss A., Schiaffino S., Lelwand L.A.;  
 RT "Comparative sequence analysis of the complete human sarcomeric myosin  
 RT heavy chain family: implications for functional diversity.";  
 RL J Mol. Biol. 290:61-75(1999).  
 RN [2]  
 RP SEQUENCE OF 1064-1939 FROM N.A.  
 RX MEDLINE=8616778; PubMed=2421254;  
 RA Saez L., Lelwand L.A.;  
 RT "Characterization of diverse forms of myosin heavy chain expressed in  
 RT adult human skeletal muscle.";  
 RL Nucleic Acids Res. 14:2951-2969(1986).  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY  
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.  
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT  
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE  
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED  
 CC SUBFRAGMENT (S2).  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



RP VARIANTS MHA/FTNS/SBS K-93; C-702; C-1165; H-1424 AND K-1841.  
RX MEDLINE-20428192; PubMed-10973259;  
RA Seri M., Cusano M., Gangarosa S., Caridi G., Bordo D., Lo Nigro C.,  
RA Ghigeri G.M., Ravazzolo R., Savino M., Del Vecchio M., d'Apolito M.,  
RA Iolascon A., Zelante L.L., Savoia A., Baldoni C.L., Norris P.,  
RA Magrini U., Bellelli S., Heath K.E., Babcock M., Glucksman M.J.,  
RA Aliprandis E., Bizzaro N., Desnick R.J., Martignetti J.A.,  
RT "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and  
RT Sebastian syndromes";  
RL Nat. Genet. 26:103-105(2000).  
RP VARIANTS MHA TLE-1155 AND LYS-1841.  
RX MEDLINE-20428193; PubMed-10973260;  
RA Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.,  
RT "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in  
RT May-Hegglin anomaly";  
RL Nat. Genet. 26:106-108(2000).  
CC -I- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,  
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND  
CC CAPING.  
CC -I- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY  
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2  
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING  
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,  
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF MAY-HEGGLIN ANOMALY  
CC (MHA), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED  
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.  
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF FECHTNER SYNDROME  
CC (FTNS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED  
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.  
CC WITH ADDITIONAL ALPORT-LIKE CLINICAL FEATURES OF SENSORINEURAL  
CC DEAFNESS, CATARACTS AND NEPHRITIS.  
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF SEBASTIAN SYNDROME  
CC (SBS), AN AUTOSOMAL DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED  
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUKOCYTE INCLUSIONS.  
CC -I- DISASE: DEFECTS IN MYH9 ARE THE CAUSE OF AN AUTOSOMAL DOMINANT  
CC FORM OF NONSYNDROMIC SENSORINEURAL DEAFNESS (DNA17) WHICH IS  
CC CHARACTERIZED BY PROGRESSIVE HEARING IMPAIRMENT AND  
CC COCHLEOSACULAR DEGENERATION.  
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
DR EMBL: 282215; CAB05105.1; -  
DR EMBL: M81105; AAA59888.1; -  
DR EMBL: M69180; AAA61765.1; -  
DR EMBL: M31013; AAA6349.1; -  
DR HSSP: P08799; 1LVK.  
DR MIM: 160775; -  
DR MIM: 153640; -  
DR MIM: 155100; -  
DR MIM: 603622; -  
DR MIM: 605249; -  
DR InterPro: IPR000048; IQ.  
DR InterPro: IPR004009; Myosin\_N.  
DR InterPro: IPR002928; Myosin\_Tail.  
DR InterPro: IPR002017; Spectrin.  
DR InterPro: IPR001609; myosin\_head.  
DR Pfam: PF00612; IQ; 1.  
DR Pfam: PF000063; myosin\_head; 1.  
DR Pfam: PF02736; myosin\_N; 1.  
DR Pfam: PF01576; myosin\_Tail; 1.  
DR PRINTS: PR00193; MYOSINHEAVY.  
DR ProDom: PD000355; myosin\_head; 1.

DR SMART; SM00015; IQ; 1.  
DR SMART; SM00242; MYSC; 1.  
DR PROSITE; PS50096; IQ; 1.  
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;  
KW Coiled coil; Alkylation; Multigene family; Disease mutation;  
KW Deafness.  
FT DOMAIN 1 778  
FT DOMAIN 779 808  
FT DOMAIN 837 1926  
FT NP\_BIND 174 181  
FT DOMAIN 654 676  
FT MOD\_RES 694 694  
FT MOD\_RES 704 704  
FT VARIANT 93 93  
FT VARIANT 702 702  
FT VARIANT 705 705  
FT VARIANT 1155 1155  
FT VARIANT 1165 1165  
FT VARIANT 1424 1424  
FT VARIANT 1841 1841  
FT CONFLICT 53 55  
FT CONFLICT 660 660  
FT CONFLICT 869 869  
FT CONFLICT 931 931  
FT CONFLICT 1240 1241  
FT CONFLICT 1350 1350  
FT CONFLICT 1764 1764  
FT CONFLICT 1771 1771  
SQ SEQUENCE 1960 AA; 226531 MW; 588F84BB8C106E6F CRC64;  
Query Match 7.0%; Score 166; DB 1; Length 1960;  
Best Local Similarity 19.6%; Pred. No. 0.37;  
Matches 88; Conservative 81; Mismatches 145; Indels 136; Gaps 15;  
QY 48 OCRIGVGRKTIINKLFEDLAOEENVIDREF-----LNKELDNVRAQLSOK 93  
DB 1073 ELKMQLAKEE--BELDAALARVEEAAQNMALKRTRELSQSLSELDLESERARNNA 1130  
QY 94 DKCKRDSQYIITDRLPTLEBRNATVVSLOALGKAE---MLCSTLKKQMKYLEQOQDET 149  
DB 1131 EKCKRDLGELTALKTLEDYTDSTAQOELRSKRQEVNIIKKTLEEAKTHEAQIOEM 1190  
QY 150 KQKQEEA-----GRLSKMKTME-----QTEILLQS-----QLPEY 180  
DB 1191 KQKHSQAVEELAEQLDQTKRKVKANLEKAKOTLENERGELANEKVLQGGDSEHRRKKV 1250  
QY 181 EEMIRDMGV-----GOSAVEQLAVYCVSLKKEYENLKEARKAGEVDAKLRKDLFFSSRK 236  
DB 1251 EAQLQELQYKFNENGERVRETLADKYKQLQVELDNVYGLLSQSSKSKLTKDPSALESQ 1310  
QY 237 QIVYSELDAQKLELKSQKQDQSAKDEIMSLKKKLTMLQETLLPVPVASEYTDRLVLESP 296  
DB 1311 Q-----DTQELQELQENRQKLSLSTRKQVEDEKN----- 1339  
QY 297 APEVNVMLKLRPSPFRDIDNATFVDTPPARSSSOHGVEYKLELKS-----HSPIDQY 352  
DB 1340 -----SFRQLE-----EEEEKKN-----LEKQIATLHAQVADM 1369  
QY 353 PKRI-----CKGPRKESQSLGQSCAGCEPDELVGAPFIYRNAILQCKQ----- 398  
DB 1370 KKKMEDSVGLTFAEYKRRKQLQKDLGLS---QRHEKVAAYKLTAKTRTRLQGLDLDL 1426  
QY 399 -PKRPSESSCKDQVVRVTFGDLGGRTKFI 427  
DB 1427 VDDHQRQACNLEKKQKKFDQLAEKTI 1456



FT VARSPLIC 2441 2704 MISSING (IN ISOFORM 4).  
 FT CONFLICT 1998 1998 G -> R (IN REF. 5).  
 SO SEQUENCE 2704 AA; 313084 MW; A7219E687A634A77 CAC64;

Query Match 7.0%; Score 166; DB 1; Length 2704;  
 Best Local Similarity 20.0%; Pred. No. 0.54;  
 Matches 97; Conservative 88; Mismatches 181; Indels 120; Gaps 19;

QY 33 CLDSFETASRRCPOQRIOVKRTIINKLFPDLAEEENVLDREFLNKELNDVRA----- 88  
 DB 1385 CRENAFLPCVITGATSCRAATG-----LQEHKQKAEEELKQVDELTAANRK 1432  
 QY 89 -----QLSKDKERKDSQVITDRLDLEERNATVVSLOQALGKAEMLCSTLKKQ 138  
 DB 1433 AEDQMRELTYELNALQLEKSSSEKALLDKLDETNTTLRLCLLELKRQDAEKGSQO 1492  
 QY 139 MKYLEQOQDET---KQAEAGRLRSKMTMEQIEL-----LLQSQLPEVEEMIRD 186  
 DB 1493 LRELGRLONTGTGKAEBAMQDASDLK-KIKRNYQLELESLNHEKGLQR---EVDRTTRA 1548  
 QY 187 MGVOGSANEVOL--AVYCVSLKKEVENKEARKASGEVADLRKDLFSSRSKLTQVYSELD 244  
 DB 1549 HAAVEKNIOHLNSQIHSEFRDEKELERLQIQKSDHLKEQFEK---SHEOLLQNIKAKE 1605  
 QY 245 QAKLELSAQKDLQASDKKEIMSLKKLTMLOETLNLPPVASETV---DRLVLESPARVE 300  
 DB 1606 NND-KIQRNLMELEKSENECKAKQVEBELTRONNEKTKIMQRQASSENTIVLEKQTIQO 1664  
 QY 301 --VNLKLRPSFRDDIDLANLTFDVPDPARPSSQHGVEKLEKSHSPIDVPPKICK 358  
 DB 1665 RCEALKTQADGFKDQL-----RSTNEH-----LHKOTKTBDQFORKI-- 1701  
 QY 359 GPRKESLSLGGSCAE---PDELYGAP-----IVRNAIILQOKQPKRRSSSC 408  
 DB 1702 -----KCLEDLAKSQNLVSEKQKCDQOQNIITQNT--KKEVRNLMAELNA 1745  
 QY 409 SKDVVATGFGDLGRTKFIQPTDVTMIRPLPVKPKTKYKQVRKYTPS-----LFO 460  
 DB 1746 SKEKRRGEGQKVOLOQAOVDELNRL-----KKVQDELHLKTIIEQMTIKKMYLFO 1796  
 QY 461 AKLDTE 466  
 DB 1797 EESGKF 1802

RESULT 26  
 USOL\_YEAST STANDARD; PRT; 1790 AA.  
 AC P25386;  
 DT 01-MAY-1992 (Rel. 22, Last Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Intracellular protein transport protein USOL.  
 GN USOL OR INT1 OR YDL058M.  
 OS Saccharomyces cerevisiae (Baker's Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180-1A;  
 RX MEDLINE=91185402; PubMed=2010462;  
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
 RA Yamasaki M.;  
 RA "A cytoskeleton-related gene, usol, is required for intracellular  
 RT protein transport in Saccharomyces cerevisiae.";  
 RL J. Cell Biol. 113:245-260(1991).  
 RN [2]  
 RP SEQUENCE OF 782-1790 FROM N.A.  
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,  
 RA Kendrick K.E.;  
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.

RN [3]  
 RP SEQUENCE OF 1-8 FROM N.A.  
 RA Bai Y., Symington L.S.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC - FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
 CC COMPLEX.  
 CC - SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC - DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED  
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
 CC - SIMILARITY: BELONGS TO THE VDP/USOL/YBL047C FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X54378; CAA38253.1; -;  
 DR EMBL: L03188; AAB00143.1; -;  
 DR EMBL: U53668; AAB66659.1; -;  
 DR PIR: A38455; A38455.  
 DR HSP: P80220; 1DIP.  
 DR SGD: S0002216; USOL.  
 DR InterPro: IPR002017; Spectrin.  
 DR TransPro: Protein transport; Golgi stack; Cytoskeleton; coiled coil.  
 FT DOMAIN 1 724 GLOBULAR HEAD.  
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).  
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).  
 FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.  
 FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).  
 FT CONFLICT 847 847 G -> E (IN REF. 2).  
 FT CONFLICT 924 924 E -> K (IN REF. 2).  
 FT CONFLICT 1253 1253 V -> I (IN REF. 2).  
 FT CONFLICT 1319 1319 I -> V (IN REF. 2).  
 FT CONFLICT 1461 1461 N -> S (IN REF. 2).  
 FT CONFLICT 1581 1581 G -> S (IN REF. 2).  
 FT CONFLICT 1600 1600 I -> V (IN REF. 2).  
 FT CONFLICT 1661 1661 R -> S (IN REF. 2).  
 FT CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).  
 SO SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CAC64;

Query Match 6.9%; Score 165.5; DB 1; Length 1790;  
 Best Local Similarity 20.9%; Pred. No. 0.36;  
 Matches 78; Conservative 75; Mismatches 129; Indels 91; Gaps 13;

QY 55 KRTIINKLFPDLA---EEENVLDREFLNKELNDVRAQLSQKDKKRSQV 102  
 DB 1368 ERKLLNGSSPTIQEYSEKWTLEDELIRLQENNELKAKEIDNTRSEL-EKVSLSND-- 1423  
 QY 103 IIDRLPTLEERNATVVSLOQAL-----GKAEMLCSTLKKMKYLEQOQDETQAOE 154  
 DB 1424 -----ELLEEKQNTISLQDEITLSYDKDKTRNDEKLLSTERDNKKRLDESLEQLRAQE 1477  
 QY 155 EAGRLRSKMTMEQIELLSQLPVEEMIRDMVGOSAVE---ALAVCVSLKKEVEN 210  
 DB 1478 SKAVEGGLKLEESSSEKAKELKSKEMMKL---ESTIESNETELKSSMETIRKDEK 1534  
 QY 211 LKEARKASGEVADKLK-----KDLFSSSKLOTYV---SELDQALEKSA 253  
 DB 1535 LEOGSADEEDIKMLQHEKSDILSRINESEKIDELKSLRIEAKSGSELETVQOELNNA 1594  
 QY 254 QKDQASDKREIMSLKKLTMLOETLNLPPVASEVDRLVLESPAPVEV---LKLRR 307  
 DB 1595 QEKIRINAEENTVYLSKL-----EDIEBELDKQAEIISNOEKEKLLSRL 1640  
 QY 308 PSFRDDIDLANLTFDVPDPARPSSQHGVEK-----LCLEKSHSPI-----Q 350

Db 1641 KELEBOLDSTOOKAKOSEERRAEVRFQVEKSQLDEKAMLETETKYNDLVNKEQAMKDE 1700  
QY 351 DVPKICKSPRKE 363  
Db 1701 DTVAKTTDSQRQE 1713

RESULT 27  
MYSR\_DIRIM  
ID\_MYSR\_DIRIM STANDARD: PRT: 848 AA.  
AC P13392;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Paramyosin (Fragment).  
OS Dirofilaria immitis (Canine heartworm).  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;  
OC Onchocercidae; Dirofilaria.  
OX NCBI\_Taxid=6287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90220759; PubMed=2325708;  
RA Limberger R.J., McReynolds L.A.;  
RT "Filarial paramyosin: cDNA sequences from Dirofilaria immitis and  
RT Onchocerca volvulus";  
RL Mol. Biochem. Parasitol. 38:271-280(1990).  
RN [2]  
RP SEQUENCE OF 18-96 FROM N.A.  
RX MEDLINE=89344126; PubMed=2527335;  
RA Granda A.G. III, Tuyen L.K., Asiklin N., Davis T.B., Philipp M.,  
RA Cohen C., McReynolds L.A.;  
RT "A lambda gt11 cDNA recombinant that encodes Dirofilaria immitis  
RT paramyosin";  
RL Mol. Biochem. Parasitol. 35:31-41(1989).  
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF  
CC MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
CC -1- SIMILARITY: HIGH, TO MYOSIN HEAVY CHAINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL: M29733; AAA28299.1; -;  
DR EMBL: J04009; AAA28300.1; -;  
DR PIR: A45548; A45548.  
DR PIR: A44972; A44972.  
DR HSSP: P80220; 1DIP.  
DR InterPro: IPR002928; Myosin\_tail.  
DR Pfam: PF01576; Myosin\_tail; 1.  
KW Coiled coil; Muscle protein; Thick filament; Myosin.  
FT NON\_TER 1  
FT DOMAIN <1 9 NONHELICAL REGION (POTENTIAL).  
FT DOMAIN 10 833 COILED COIL (POTENTIAL).  
FT DOMAIN 834 848 NONHELICAL REGION (POTENTIAL).  
FT DISULFID 105 105 INTERCHAIN (POTENTIAL).  
FT DISULFID 594 594 INTERCHAIN (POTENTIAL).  
FT NON\_TER 848 848  
SQ SEQUENCE 848 AA; 98008 MW; BF32B7120157FA50 CRC64;

Query Match 6.9%; Score 165; DB 1; Length 848;  
Best Local Similarity 19.2%; Pred. No. 0.16;  
Matches 102; Conservative 90; Mismatches 161; Indels 178; Gaps 19;

QY 59 INKLFPLAD-----EENVLDREF--LNKELDN---VRAQLSQKKEK----- 98  
Db 171 LNRHVNDLAQORORLAQENNDLKEIHQKYOVDLNDLQHVYQYLAQOLEFAARRLEDAERE 230

QY 99 -----DSQVITDRLDRTLEERNATVVSLOQALGA----- 128  
Db 231 RSQLOAQLOHVOLELDSVRYTALDEESARAEKRLALATETIOMKSKFDVAEVALHHE 290  
QY 129 --EMICSTLKRMKMYLEBOOQ---DETQAOEAEAGRLSKMKT----- 166  
Db 291 VEDLRKKMLQKQAEYEQIEIMLOKISOLEKAKSRLQSEVEVLIVDEKAQNTAILERA 350  
QY 167 -EQIELLOSQLPEVEEIRDMVGQ-----SAVEQLAVYCVSLKE 207  
Db 351 KEOLEKTVNELKVRIDELTYLEAQAALAELOKMKNLKAEIOKALARENKKL 410  
QY 208 YENLKEARKASGEVADRLK-DLFSSR-----SKLOTVSEL 243  
Db 411 QDDLHEAKELADANKRLHELDLENARLAGIRELOALAKESEARDAERARALAE 470  
QY 244 DQAKLELSAQKDSADKEITMSLKKRLTMLQETLNPVASETVDR- --VLESPAPVE 300  
Db 471 QQLRTM---ERRLOKEKEEAEALRKNMQF-----EIDRLTALADAEARMK 514  
QY 301 VNLKLRPSFRDQI-DLNATFD-----VTPPARPSSOHGYEKEKLEKSHSPIDVP 353  
Db 515 AEISRLKKKQAEIAELMTVDNLRANIEAQTKIKQSEQLKILQASLEDTOROLOTL 574  
QY 354 KKICKGPRKESQLSLGGSCAGEPDELVGAFPIFVRNALIGOKOPRRPSSSCSDVY 413  
Db 575 DQYLAQRKYSALSAELEC-----KALDNALTAARKQAEIDLEANA----- 617  
QY 414 RTGFGLGGRTRFTPTDVTVMRPLPVKPKTKVRQVRKVTSPSLFOAKLD 464  
Db 618 -----RTDVLVSIN-----NNLTAKIN--KLETELSTAQADLD 648

RESULT 28  
RAD50\_ARCFU  
ID\_RAD50\_ARCFU STANDARD: PRT: 886 AA.  
AC Q29230;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE DNA double-strand break repair rad50 Arpase.  
GN RAD50 OR AFI032.  
OS Archaeoglobus fulgidus.  
CC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
CC Archaeoglobus.  
CC NCBI\_Taxid=2234;  
OX  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kierlavage A.R., Graham D.E., Kypides N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirchness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artlach P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).  
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The  
CC rad50/mre11 complex possesses single-strand endonuclease activity  
CC and ATP-dependent double-strand-specific exonuclease activity.  
CC Rad50 provides an ATP-dependent control of mre11 by unwinding  
CC and/or repositioning DNA ends into the mre11 active site (By  
CC similarity).  
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.









DR SMART; SM00180; EGF\_Lam; 10.  
 DR SMART; SM00001; EGF\_Like; 1.  
 DR SMART; SM00281; Lam; 1.  
 DR SMART; SM00136; Lam; 1.  
 DR PROSITE; PS00022; EGF\_L; 8.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01248; LAMININ\_Type\_EGF; 11.  
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 FT SIGNAL 1 33 POTENTIAL.  
 FT CHAIN 34 1639 LAMININ GAMMA-1 CHAIN.  
 FT DOMAIN 34 298 LAMININ N-TERMINAL (DOMAIN VI).  
 FT DOMAIN 299 358 LAMININ EGF-Like 1.  
 FT DOMAIN 359 413 LAMININ EGF-Like 2.  
 FT DOMAIN 414 460 LAMININ EGF-Like 3.  
 FT DOMAIN 461 513 LAMININ EGF-Like 4.  
 FT DOMAIN 514 523 LAMININ EGF-Like 5 (N-TERMINAL).  
 FT DOMAIN 524 709 LAMININ DOMAIN IV.  
 FT DOMAIN 710 743 LAMININ EGF-Like 6.  
 FT DOMAIN 744 792 LAMININ EGF-Like 7 (C-TERMINAL).  
 FT DOMAIN 793 846 LAMININ EGF-Like 8.  
 FT DOMAIN 847 901 LAMININ EGF-Like 9.  
 FT DOMAIN 902 955 LAMININ EGF-Like 10.  
 FT DOMAIN 956 1003 LAMININ EGF-Like 11.  
 FT DOMAIN 1004 1049 LAMININ EGF-Like 1.  
 FT DOMAIN 1050 1609 LAMININ EGF-Like 10.  
 FT DOMAIN 1087 1109 COILED COIL (POTENTIAL).  
 FT DOMAIN 1144 1247 COILED COIL (POTENTIAL).  
 FT DOMAIN 1306 1627 COILED COIL (POTENTIAL).  
 FT DOMAIN 1299 308 BY SIMILARITY.  
 FT DISULFID 301 322 BY SIMILARITY.  
 FT DISULFID 324 333 BY SIMILARITY.  
 FT DISULFID 336 358 BY SIMILARITY.  
 FT DISULFID 359 366 BY SIMILARITY.  
 FT DISULFID 361 384 BY SIMILARITY.  
 FT DISULFID 387 396 BY SIMILARITY.  
 FT DISULFID 399 411 BY SIMILARITY.  
 FT DISULFID 414 426 BY SIMILARITY.  
 FT DISULFID 416 432 BY SIMILARITY.  
 FT DISULFID 434 443 BY SIMILARITY.  
 FT DISULFID 446 458 BY SIMILARITY.  
 FT DISULFID 461 475 BY SIMILARITY.  
 FT DISULFID 475 482 BY SIMILARITY.  
 FT DISULFID 484 493 BY SIMILARITY.  
 FT DISULFID 496 511 BY SIMILARITY.  
 FT DISULFID 511 523 BY SIMILARITY.  
 FT DISULFID 744 753 BY SIMILARITY.  
 FT DISULFID 746 760 BY SIMILARITY.  
 FT DISULFID 762 771 BY SIMILARITY.  
 FT DISULFID 774 790 BY SIMILARITY.  
 FT DISULFID 793 801 BY SIMILARITY.  
 FT DISULFID 795 811 BY SIMILARITY.  
 FT DISULFID 814 823 BY SIMILARITY.  
 FT DISULFID 826 844 BY SIMILARITY.  
 FT DISULFID 847 861 BY SIMILARITY.  
 FT DISULFID 868 880 BY SIMILARITY.  
 FT DISULFID 871 880 BY SIMILARITY.  
 FT DISULFID 883 899 BY SIMILARITY.  
 FT DISULFID 902 919 BY SIMILARITY.  
 FT DISULFID 926 936 BY SIMILARITY.  
 FT DISULFID 937 953 BY SIMILARITY.  
 FT DISULFID 940 953 BY SIMILARITY.  
 FT DISULFID 956 968 BY SIMILARITY.  
 FT DISULFID 958 975 BY SIMILARITY.  
 FT DISULFID 977 986 BY SIMILARITY.  
 FT DISULFID 989 1001 BY SIMILARITY.  
 FT DISULFID 1004 1016 BY SIMILARITY.  
 FT DISULFID 1006 1022 BY SIMILARITY.  
 FT DISULFID 1024 1033 BY SIMILARITY.  
 FT DISULFID 1036 1047 BY SIMILARITY.  
 FT DISULFID 1050 1050 BY SIMILARITY.  
 FT DISULFID 1053 1053 BY SIMILARITY.  
 FT DISULFID 1631 1631 BY SIMILARITY.  
 FT CARBOHYD 147 147 (POTENTIAL).

FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 862 862 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 965 965 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 Query Match 6.9%; Score 163.5; DB 1; Length 1639;  
 Best Local Similarity 21.7%; Pred. No. 0.4;  
 Matches 117; Conservative 72; Mismatches 220; Indels 129; Gaps 21;  
 QY 7 CTISDFHDSRDVAALHCHGTFHLOCLIOSFETAPSRTPQCRIOYKRTIINKLEFPL 66  
 DB 1047 CIDPCDYNLVQDAADLHRAKLFNLSQTLDEIARTPTVNDDEE--AKLVAQEKYAVL 1103  
 QY 67 AOEEN-----VLDRFLKNELDNVRAOLSOKEKRPDSQYITLRLTEERN 115  
 DB 1104 AQDARDSDGSGOGGYAAVEYID--LHKHLDSYRELHVSADKFOADANEIDRAR-----QN 1156  
 QY 116 ATVV-----SLQALG----KAEMLCSTLKKOMKYLEQOODETKQAQAEAGRLRS 161  
 DB 1157 YTIIDQITENAKKELQOALDLNDEGAQALARAKEYSVEF--GOOSEQISDISREARALAD 1215  
 QY 162 KMKTMQDEL-----LLQSLPVEVEIMRMGVSQSAVQLAYVCV 202  
 DB 1216 KLESEAFDLKNAKADAKAVEKAHQAKSAIDQLKIGTELRSVGLTSHVKQ----- 1269  
 QY 203 SLKKEYENLKEARKASGEVADKLRLKDLFSSRSKLTQTYVELDQAKLELKSADKLOSDAK 262  
 DB 1270 SLGTVGTSTKFALEKANEYVD-----TALTILNDV--NQTOPEIDISOLKDAVANE 1321  
 QY 263 EIMSLKRLKLTMLQFTLMLPVASETVDRVLSESPAVYVNLKLRPSFRDIDL---NA 318  
 DB 1322 RADELALKQITELNSNGELPADFETQEL-----TEALKRAEQOOLDIELLERAKA 1374  
 QY 319 TFDVDTPPARPSSS-----QHCYEEKLC-----LEKSHSPIDYPPKRICGPRRESQ 365  
 DB 1375 AHDKATRAVEGDDTLEKANNTYEKLAFQSDVQSSSESEKALQTVP-----NIEKIQ 1429  
 QY 366 LSLGQSCAGEPDELVGAPPIEFVRNALIGOKOP-----KRPRESS 407  
 DB 1430 ---NAESLIQAEALDGA---NKNANEAKKNAQAEQLKAYADASDAELIRKANETK 1482  
 QY 408 CSKDVRFTGDLGGRFTQPTDTVMIRPLPVKPKTKVQRVRYKTPSLGFAKLD 465  
 DB 1483 VAARNLREADDQLNHRVKLTE-MDIFKLEESSTFDMLVYDARK-----KVGAQAKDT 1534  
 RESULT 33  
 ALM1\_SCHPO STANDARD; PRT; 1727 AA.  
 ID ALM1\_SCHPO  
 AC Q9UTR5; O13313; Q9UTR8;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Abnormal long morphology protein 1 (Sp8).  
 GN ALM1 OR SPAC1486.04C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=972;  
 RA McDougall R.C., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;  
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RC SEQUENCE OF 495-1727 FROM N.A., AND CHARACTERIZATION.  
 RP STRAIN=972;  
 RX MEDLINE=20123449; PubMed=10660053;  
 RA Jimenez M., Petit T., Gancedo C., Goday C.;  
 RT "The alm1+ gene from Schizosaccharomyces pombe encodes a coiled-coil  
 protein that associates with the medial region during mitosis.";  
 RL Mol. Gen. Genet. 262:921-930(2000).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

RN [3]  
 RP SEQUENCE OF 644-834 FROM N.A.  
 RC STRAIN-968 H90;  
 RA Ding D.;  
 RT "Generation and analysis of GFP-gene fusion library of fission yeast."  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: AFTER THE ONSET OF MITOSIS, AT MID- TO LATE ANAPHASE, CO-LOCALIZES WITH THE MEDIAL ACTIN RING. MAY PLAY A ROLE IN CYTOKINESIS.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL13357; CAB62414.1; -  
 DR EMBL: AF010473; AAB65416.1; ALT\_INIT.  
 DR EMBL: AB028012; BAA87316.1; -  
 KW Coiled coil.  
 FT DOMAIN 57 361 COILED COIL (POTENTIAL).  
 FT DOMAIN 443 463 COILED COIL (POTENTIAL).  
 FT DOMAIN 542 740 COILED COIL (POTENTIAL).  
 FT DOMAIN 804 1106 COILED COIL (POTENTIAL).  
 FT DOMAIN 1223 1427 COILED COIL (POTENTIAL).  
 FT DOMAIN 1497 1555 COILED COIL (POTENTIAL).  
 FT DOMAIN 1601 1664 COILED COIL (POTENTIAL).  
 SQ SEQUENCE 1727 AA; 197858 MW; F820BF8D9C132644 CRC64;

Query Match 6.9%; Score 163.5; DB 1; Length 1727;  
 Best Local Similarity 22.8%; Pred. No. 0.43;  
 Matches 97; Conservative 60; Mismatches 132; Indels 137; Gaps 17;

QY 30 HLCGLISFETAPSRCTPCORIO---VGKRTIINKL---PFDLAQEBENV-----LDRE 77  
 DB 1301 HNGEITLILSNSTSTSSDASRLKNELYSKENLIELNQELGHLSSELEYKSKSEDIENE 1360  
 QY 78 FLKN-----ELDN-----VRAQLSQDKERKRSQVIIDT 106  
 DB 1361 RAQMQSKIEQLEKNTKLAAMRTKYGVYVKSLEKHNRQQLSQKSTSELEAKVACHQ 1420  
 QY 107 LKRTLEERNAT-----VLSQALGKAEMLCSTLKKQMYL----- 142  
 DB 1421 LNEQLNKPASATPTATQSEPTVSLEEFNSTKELSTORLSTEDIMLNTTKEELEKVR 1480  
 QY 143 -----EODDETKQAOEAGRLRSKM-KTMEQIELL-----LOSOLP 178  
 DB 1481 QNSNKSGETSKDTPIPEDEEKKKVMQOEVLRLSRKAELQNLKRONOVLOQOVK 1540  
 QY 179 EVEDMIRMGVGSQAVQOLAVYCVSLKREYNLEARK-----ASG 219  
 DB 1541 ALQETV-----VSSEASASVHADT-KDLENLKTTEEMLSVTFQVIFNESISDFSTIA 1594  
 QY 220 EVADKLKDLFSSRSKQTYISELDQAKLEKSKQKDLQSDAKET-MSLKKLTLMLQETL 278  
 DB 1595 DFTTFVQKEWKKREIILQK-----DVEEOVAVSHQKQDLNIRKLEEMRNKLKLSMLRK- 1647  
 QY 279 NLPVASETDVRLVLESPAPVEVNLKLRPSFRDDIDLNA-----TFPVDR--P 325  
 DB 1648 NLAVRYALEDSKKKSDPAIILSEASKNTDSNKSNSVPAQVKEKKLIAKTAHSDVTNSP 1707  
 QY 326 PARPSS 331  
 DB 1708 PKRSSS 1713  
 RESULT 34  
 MYHA\_HUMAN

ID MYHA\_HUMAN STANDARD; PRT; 1976 AA.  
 AC P35580;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, nonmuscle type B (Cellular myosin heavy chain, type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).  
 GN MYH10.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96025307; PubMed=7499478;  
 RA Phillips C.L., Yamakawa K., Adelstein R.S.;  
 RT "Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and analysis of human tissues with isoform-specific antibodies."  
 RL J. Muscle Res. Cell Motil. 16:379-389(1995).  
 RN [2]  
 RP SEQUENCE OF 63-722 FROM N.A.  
 RX MEDLINE=91316803; PubMed=1860190;  
 RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K., Gula D., Adelstein R.S., Weir L.;  
 RT "Human nonmuscle myosin heavy chains are encoded by two genes located on different chromosomes."  
 RL Circ. Res. 69:530-539(1991).  
 CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS, CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND CAPING.  
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M69181; AAA99177.1; -  
 DR PIR: B61231; B61231.  
 DR HSSP: P08799; ILVK.  
 DR MIM: 160776; -  
 DR InterPro: IPR000048; IQ.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR002017; Spectrin.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF02736; Myosin\_N; 1.  
 DR Pfam: PF01576; Myosin\_tail; 1.  
 DR PRINTS: PR00193; MYOSINHEAVY.  
 DR ProDom: PD000355; myosin\_head; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00242; MYSC; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;  
 KW Coiled coil; Alkylation; Multigene family.  
 FT DOMAIN 1 785 MYOSIN HEAD-LIKE.  
 FT DOMAIN 786 815 IQ.  
 FT DOMAIN 845 1976 COILED COIL (POTENTIAL).  
 FT NP\_BIND 178 185 ATP (POTENTIAL).  
 FT MOD\_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).  
 FT MOD\_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).

SQ SEQUENCE 1976 AA; 228938 MW; B2B87FF35EA124F CRC64;  
 Query Match 6.9%; Score 163.5; DB 1; Length 1976;  
 Best Local Similarity 22.78; Pred. No. 0.5;  
 Matches 113; Conservative 77; Mismatches 152; Indels 155; Gaps 21;  
 QY 35 IQSFTAPSPFCPOCRIOVGKRTITNKLFPDLAOEENVLDRP--LKNELDNVRAQLSQ 92  
 DB 1387 IESLEEKKKLLKDAE-ALSGRLEEKALAYDKLEKTKNRLOQELDDLTVDLDRQVASN 1445  
 QY 93 -KDEKRDQOVIIDTL-----RDTLF---ERNATVVSLOQALGKA----- 128  
 DB 1446 LKKKKKKFDOILAEKSSISARYAEERDAEAREKETKALSLARALEEAKKEFERQ 1505  
 QY 129 -----EMCST-----LKKOMKYLEGOODETK-----OAOEAG-RL 159  
 DB 1506 NKQLRADMEDLMSKDKVGNVHELEKSKRALBOOVEMRQLELEDELDQATEDAKRL 1565  
 QY 160 RSKNKT-----EQLLELLLOSLPEVEEMIRMGVGSQAVQLAVCVSL 204  
 DB 1566 EVNQAKMAKQAFERDLQTRDEONEKKRLILIKQVRELEAELED-----ERQRLAVASK 1619  
 QY 205 KKEVENKE-----ARKASGEVADKLR-----DLFS-----SR 233  
 DB 1620 KKMEDIDKLEAQIEAANKARDEVYIKQLRLQAOQMDYQRELEBARASRDEIFAQSESE 1679  
 QY 234 SKLOTVSELDQAKLEKSQKDSADKEIMSLKKKIT-----MLQETLPLPVAS 285  
 DB 1680 KKLKSLLEHETIQLOEELASSERARRHAEQERDELADETINSAGSKSLDRERLEARIA 1739  
 QY 286 ETYVRLVLESPAPVEVNLKLRPSFRDDIDLVDTPPARSSSOHGYYERLCLEKS 345  
 DB 1740 QLEELEEE-----QSMELLNDRFR-----KTLQVDTLNA-----ELAERS 1778  
 QY 346 HSPILQDVPKTKCKPKRESQSLGQSCAGEPDELYGAF-----PIFYANALGQ-K 397  
 DB 1779 AAKSDNARQOLEKRONELKRLKLT-----QELGAVKSKFKATISALEAKIQGLEE 1828  
 QY 398 QPKRPRESSCKDVR 414  
 DB 1829 QLEQAKERAKANKVR 1845  
 RESULT 35  
 ID DESP\_HUMAN STANDARD; PRT; 2871 AA.  
 AC P15924; Q14189; O75993; Q9UHN4;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Desmoplakin (DP) (250/210 kDa paraneoplastic pemphigus antigen).  
 GN DSP.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SPOUNCE FROM N.A. (ISOFORM DPI).  
 RC TISSUE-Forceskin;  
 RX MEDLINE=92115697; PubMed=1731325;  
 RA Virata M.L.A., Wagner R.M., Parry D.A.D., Green K.J.;  
 RT "Molecular structure of the human desmoplakin I and II amino  
 RT terminus";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:544-548(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM DPI).  
 RA Phillips S.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SPOUNCE OF 1120-2871 FROM N.A. (ISOFORM DPI).  
 RC TISSUE-Forceskin;  
 RX MEDLINE=90153880; PubMed=1689290;

RA Green K.J., Parry D.A.D., Steinert P.M., Virata M.L.A., Wagner R.M.,  
 RA Angst B.D., Nilles L.A.;  
 RT "Structure of the human desmoplakins. Implications for function in  
 RT the desmosomal plaque";  
 RL J. Biol. Chem. 265:2603-2612(1990).  
 RN [4]  
 RP ERRATUM.  
 RX MEDLINE=90361712; PubMed=2391353;  
 RA Green K.J., Parry D.A.D., Steinert P.M., Virata M.L.A., Wagner R.M.,  
 RA Angst B.D., Nilles L.A.;  
 RL J. Biol. Chem. 265:11406-11407(1990).  
 RN [5]  
 RP SEQUENCE OF 2854-2871 FROM N.A.  
 RC TISSUE-Skin;  
 RX MEDLINE=20062965; PubMed=10594734;  
 RA Whitlock N.V., Ashton G.H., Dopling-Hepenstal P.J., Gratian M.J.,  
 RA Keane F.M., Eady R.A.J., McGrath J.A.;  
 RT "Striate palmoplantar keratoderma resulting from desmoplakin  
 RT haploinsufficiency";  
 RL J. Invest. Dermatol. 113:940-946(1999).  
 RN [6]  
 RP CHARACTERIZATION.  
 RX MEDLINE=98012209; PubMed=9348293;  
 RA Kowalczyk A.P., Bornslaeger E.A., Borgwardt J.E., Palke H.L.,  
 RA Dhaliwal A.S., Corcoran C.M., Denning M.F., Green K.J.;  
 RT "The amino-terminal domain of desmoplakin binds to plakoglobin and  
 RT clusters desmosomal cadherin-plakoglobin complexes";  
 RL J. Cell Biol. 139:773-784(1997).  
 CC - FUNCTION: MAJOR HIGH MOLECULAR WEIGHT PROTEIN OF DESMOSOMES.  
 CC INVOLVED IN THE ORGANIZATION OF THE DESMOSOMAL CADHERIN-  
 CC PLAKOGLIBIN COMPLEXES INTO DISCRETE PLASMA MEMBRANE DOMAINS AND IN  
 CC THE ANCHORING OF INTERMEDIATE FILAMENTS TO THE DESMOSOMES.  
 CC - SUBUNIT: HOMODIMER.  
 CC - SUBCELLULAR LOCATION: INNERMOST PORTION OF THE DESMOSOMAL PLAQUE.  
 CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS: DPI/DPI (SHOWN HERE) AND  
 CC DPI/DP2; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC - TISSUE SPECIFICITY: DPI IS APPARENTLY AN OBLIGATE CONSTITUENT OF  
 CC ALL DESMOSOMES; DPI RESIDE PREDOMINANTLY IN TISSUES AND CELLS OF  
 CC STRATIFIED ORIGIN.  
 CC - DOMAIN: THE N-TERMINAL REGION IS REQUIRED FOR LOCALIZATION TO THE  
 CC DESMOSOMAL PLAQUE AND INTERACTS WITH THE N-TERMINAL REGION OF  
 CC PLAKOGLIBIN 1. THE C-TERMINAL REGION INTERACTS WITH INTERMEDIATE  
 CC FILAMENTS.  
 CC - DISEASE: DEFECTS IN DSP ARE A CAUSE OF STRIATE PALMOPLANTAR  
 CC KERATODERMA II (PKR2, KPP2 OR SPPK2), CHARACTERIZED BY SKIN  
 CC THICKENING IN THE PALMS (LINEAR PATTERN) AND THE SOLES (ISLAND-  
 CC LIKE PATTERN) AND FLEXOR ASPECT OF THE FINGERS; AND RARELY BY  
 CC ABNORMALITIES OF THE NAILS, THE TEETH AND THE HAIR.  
 CC - SIMILARITY: CONTAINS 17 PLECTRIN REPEATS.  
 CC - SIMILARITY: CONTAINS 2 SPECTRIN REPEATS.  
 CC - SIMILARITY: BELONGS TO THE PLAKIN OR CYTOLINKER FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: W77830; AAA85135.1; -  
 CC EMBL: AL031058; CAI19927.1; -  
 CC EMBL: J05211; AAA35766.1; -  
 CC EMBL: AF139065; AAP19785.1; -  
 CC PIR: A35536; A35536.  
 CC HSSP: P01100; IFOS.  
 CC MIM: 125647; -  
 CC InterPro: IPR001101; Plectin\_repeat.  
 CC InterPro: IPR002017; Spectrin.  
 CC Pfam: PF00681; Plectin\_repeat; 20.  
 CC SMART: SM00250; PLEC; 17.  
 CC SMART: SM00150; SPEC; 1.  
 KW Repeat; Coiled coil; Phosphorylation; Cytoskeleton;

```

KW Structural protein; Alternative splicing.
FT DOMAIN 1 1056
FT 1057 1945 CENTRAL FIBROUS ROD DOMAIN.
FT 1946 2871 GLOBULAR 2.
FT 347 447 SPECTRIN 1.
FT 858 952 SPECTRIN 2.
FT 1018 1945 COILED COIL (POTENTIAL).
FT 2009 2045 PLECTIN 1.
FT 2046 2083 PLECTIN 2.
FT 2084 2121 PLECTIN 3.
FT 2122 2159 PLECTIN 4.
FT 2163 2197 PLECTIN 5.
FT 2198 2233 PLECTIN 6.
FT 2251 2288 PLECTIN 7.
FT 2289 2326 PLECTIN 8.
FT 2327 2364 PLECTIN 9.
FT 2365 2402 PLECTIN 10.
FT 2406 2440 PLECTIN 11.
FT 2456 2493 PLECTIN 12.
FT 2507 2544 PLECTIN 13.
FT 2610 2647 PLECTIN 14.
FT 2648 2685 PLECTIN 15.
FT 2724 2761 PLECTIN 16.
FT 2762 2799 PLECTIN 17.
FT 2824 2847 6 X 4 AA TANDEM REPEATS OF G-S-R-[SR].
FT DOMAIN 1 INTERACTS WITH PLAKOPHILIN 1 AND JUNCTION
FT 1 PLAKOGLOBIN.
FT MOD_RES 2849 2849 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT VARSLIC 1195 1794 MISSING (IN ISOFORM DELT).
FT CONFLICT 905 905 R -> A (IN REF. 2).
FT CONFLICT 1120 1120 D -> R (IN REF. 3).
FT CONFLICT 2687 2688 SV -> RL (IN REF. 2).
SQ SEQUENCE 2871 AA; 331771 MW; 5639587CD469087 CMC64;

```

```

Query Match 6.9%; Score 163.5; DB 1; Length 2871;
Best Local Similarity 20.3%; Pred. No. 0.77;
Matches 96; Conservative 96; Mismatches 156; Indels 125; Gaps 22;

```

```

QY 15 DHSRDVAHICGHTFHQCLIQSFETAPSRTCQCRIQVGKRTIINKLFPDLQAEENVL 74
DB 1187 EYENELKAVNNHNEESNLRNKET-----EINIKTKT-----KEISQKQKD-- 1230
QY 75 DREFLKNELDNVAQLSQDKERKDSQVIID--TLRDTLEBNATVVSLOALGKAEMLC 132
DB 1231 DSKNLRNQD-----RLSRENRDLKDELIVRLDSTLQATEQRRAEENALQOKAGSGEIMQ 1286
QY 133 ST--LKKQKTYLEQODET---KQAEAGR-LRSKMKTMEDIELLQ-----S 175
DB 1287 KKHLEIELKQVMQQRSEDAHRHQSLEFAKTIDQKNKEIERLKAFFEAKRMEYEN 1346
QY 176 QLEEV-----EEMI-----RDMGVQSAVEQLAVY---CVSLKREYENLKREARASG 219
DB 1347 ELSLVNRRNYDEEITSLKNQFETETINIKTKTHQTMQKEEDTSGYRQIDNLTRENSLS 1406
QY 220 EVADKLKDLFSSRSKQTYSELQDAK-----LELKA-----Q 254
DB 1407 EETIKRLKNTLQTTENLRARVEDIDQOKATGSEVSQKQKQLEVLRYQVQTMRTSESVRYK 1466
QY 255 KDIQADKETSLSLKKLTLMQETLNPPVASETVYDRLVLESPAPVEVNLKLRPSPDDI 314
DB 1467 QSDDDAKTIDQKNKEIERKQID-----KETNDRKCLE-----DENARLQRYQY--- 1512
QY 315 DLNATFVDTPPARPSSSGHY-----YEKLCLEKSHSPIDQVPPKIKKGPRKESOLS 367
DB 1513 DLOKANSATETINKLVQDEQLTRLRIDYERVSQERTVKD--QDITR--FQNSLKEQLQ 1569
QY 368 LGGSGCAGEPDEDELVGAFPIFVRNALIGOKQPKRPSSESCSKDVYRTGFDGL 420
DB 1570 -----KQKVEEEL-----NRKRTASEDSCKRKKLDEELEGM 1601

```

RESULT 36

```

MYH7_HUMAN
ID MYH7_HUMAN STANDARD: PRT; 1935 AA.
AC P12883; O14904; O16579;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, cardiac muscle beta isoform (MYHC-beta).
GN MYH7 OR MYCIB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91065634; PubMed=2249844;
RA Jaenicke T., Diederich K.W., Haas W., Schleich J., Lichter P.,
RA Florid M., Bach A., Vosberg H.P.;
RT "The complete sequence of the human beta-myosin heavy chain gene and
RL Genomics 8:194-206(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90301496; PubMed=2362820;
RA Liew C.-C., Sole M.J., Yamauchi-Takihara K., Kellam B.,
RA Anderson D.H., Lin L., Liew J.;
RT "Complete sequence and organization of the human cardiac beta-myosin
RL heavy chain gene.";
RN [3]
RP SEQUENCE OF 1-115 FROM N.A.
RX MEDLINE=89264452; PubMed=2726733;
RA Yamauchi-Takihara K., Sole M.J., Liew J., Ing D., Liew C.-C.;
RT "Characterization of human cardiac myosin heavy chain genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:3504-3508(1989).
RN [4]
RP ERRATUM.
RA Yamauchi-Takihara K., Sole M.J., Liew J., Ing D., Liew C.-C.;
RL Proc. Natl. Acad. Sci. U.S.A. 86:7416-7417(1989).
RN [5]
RP SEQUENCE OF 1310-1935 FROM N.A.
RX MEDLINE=86176778; PubMed=2421254;
RA Saez L., Leinwand L.A.;
RT "Characterization of diverse forms of myosin heavy chain expressed in
RL adult human skeletal muscle.";
RL Nucleic Acids Res. 14:2951-2969(1986).
RP REVISIONS.
RA Leinwand L.A.;
RL Submitted (MAR-1988) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 1410-1935 FROM N.A.
RX MEDLINE=88299163; PubMed=2969919;
RA Kurabayashi M., Tsuchimochi H., Komuro I., Takaku F., Yazaki Y.;
RT "Molecular cloning and characterization of human cardiac alpha- and
RT beta-form myosin heavy chain complementary DNA clones. Regulation of
RT expression during development and pressure overload in human
RL atrium.";
RL J. Clin. Invest. 82:524-531(1988).
RN [8]
RP SEQUENCE OF 785-1935 FROM N.A.
RX TISSUE=Skeletal muscle;
RL MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goede H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RL myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [9]
RP SEQUENCE OF 1393-1935 FROM N.A.
RX MEDLINE=87192738; PubMed=3032769;
RA Jandreski M.A., Liew C.-C.;
RT "Construction of a human ventricular cDNA library and
RT characterization of a beta myosin heavy chain cDNA clone.";

```

RL Hum. Genet. 76:47-53(1987).  
 RN [10]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE=96039076; PubMed=8533830;  
 RA Arai S., Matsuoaka R., Hirayama K., Sukurai H., Tamura M., Ozawa T.,  
 RA Kimura M., Imamura S.-T., Fututani Y., Joh-O K., Kawana M., Takao A.,  
 RA Hosoda S., Momma K.;  
 RT "Missense mutation of the beta-cardiac myosin heavy-chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Am. J. Med. Genet. 58:267-276(1995).  
 RN [11]  
 RP VARIANTS CMH1 GLU-256 AND ARG-741.  
 RX MEDLINE=93248216; PubMed=8483915;  
 RA Fananapazir L., Dalakas M.C., Cyran F., Cohn G., Epstein N.D.;  
 RT "Missense mutations in the beta-myosin heavy-chain gene cause central  
 RT core disease in hypertrophic cardiomyopathy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3993-3997(1993).  
 RN [12]  
 RP VARIANT CMH1 GLN-403.  
 RX MEDLINE=90367131; PubMed=1975517;  
 RA Geisterfer-Loefer A.A.T., Kass S., Tanigawa G., Vosberg H.-P.,  
 RA McKenna W., Seidman C.E., Seidman J.G.;  
 RT "A molecular basis for familial hypertrophic cardiomyopathy: a beta  
 RT cardiac myosin heavy chain gene missense mutation.";  
 RL Cell 62:999-1006(1990).  
 RN [13]  
 RP VARIANTS CMH1 GLN-249; GLN-403; CYS-453 AND MET-606.  
 RX MEDLINE=92204193; PubMed=1552912;  
 RA Watkins H., Rosenzweig A., Hwang D.S., Levi T., McKenna W.,  
 RA Seidman C.E., Seidman J.G.;  
 RT "Characteristics and prognostic implications of myosin missense  
 RT mutations in familial hypertrophic cardiomyopathy.";  
 RL New Engl. J. Med. 326:1108-1114(1992).  
 RN [14]  
 RP VARIANTS CMH1 GLN-403; CYS-453; ARG-584 AND MET-606.  
 RX MEDLINE=94070863; PubMed=8250038;  
 RA Watkins H., Thierfelder L., Anan R., Jarcho J., Matsumori A.,  
 RA McKenna W., Seidman J.G., Seidman C.E.;  
 RT "Independent origin of identical beta cardiac myosin heavy-chain  
 RT mutations in hypertrophic cardiomyopathy.";  
 RL Am. J. Hum. Genet. 53:1180-1185(1993).  
 RN [15]  
 RP VARIANTS CMH1 GLN-403 AND VAL-908.  
 RX MEDLINE=92346810; PubMed=1638703;  
 RA Epstein N.D., Cohn G.M., Cyran F., Fananapazir L.;  
 RT "Differences in clinical expression of hypertrophic cardiomyopathy  
 RT associated with two distinct mutations in the beta-myosin heavy chain  
 RT gene. A 908Leu->Val mutation and a 403Arg->Gln mutation.";  
 RL Circulation 86:345-352(1992).  
 RN [16]  
 RP VARIANTS CMH1 LEU-403 AND TRP-403.  
 RX MEDLINE=94075629; PubMed=8254035;  
 RA Dussé E., Komajda M., Feltner L., Dubourg O., Dufour C., Carrier L.,  
 RA Wisniewsky C., Bercevic J., Hengstenberg C., Al-Mahdawi S.;  
 RT "Familial hypertrophic cardiomyopathy. Microsatellite haplotyping and  
 RT identification of a hot spot for mutations in the beta-myosin heavy  
 RT chain gene.";  
 RL J. Clin. Invest. 92:2807-2813(1993).  
 RN [17]  
 RP VARIANT CMH1 TRP-403.  
 RX MEDLINE=94093568; PubMed=8268932;  
 RA Moollan J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a new missense mutation at Arg403, a Cpg mutation  
 RT hotspot, in exon 13 of the beta-myosin heavy chain gene in  
 RT hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 2:1731-1732(1993).  
 RN [18]  
 RP VARIANT CMH1 ASN-615.  
 RX MEDLINE=93038688; PubMed=1417858;  
 RA Nishi H., Kimura A., Harada H., Toshihima H., Sasazuki T.;  
 RT "Novel missense mutation in cardiac beta myosin heavy chain gene  
 RT found in a Japanese patient with hypertrophic cardiomyopathy.";  
 RL Biochem. Biophys. Res. Commun. 188:379-387(1992).

RN [19]  
 RP VARIANT CMH1 GLY-778.  
 RX MEDLINE=93343938; PubMed=8343162;  
 RA Harada H., Kimura A., Nishi H., Sasazuki T., Toshihima H.;  
 RT "A missense mutation of cardiac beta-myosin heavy chain gene linked  
 RT to familial hypertrophic cardiomyopathy in affected Japanese  
 RT families.";  
 RL Biochem. Biophys. Res. Commun. 194:791-798(1993).  
 RN [20]  
 RP VARIANT CMH1 VAL-908.  
 RX MEDLINE=93168485; PubMed=8435239;  
 RA Al-Mandawi S., Chamberlain S., Cleland J., Nihoyannopoulos P.,  
 RA Gilligan D., French J., Choudhury L., Williamson R., Oakley C.;  
 RT "Identification of a mutation in the beta cardiac myosin heavy chain  
 RT gene in a family with hypertrophic cardiomyopathy.";  
 RL Br. Heart J. 69:136-141(1993).  
 RN [21]  
 RP VARIANT CMH1 TRP-719.  
 RX MEDLINE=95179132; PubMed=7874131;  
 RA Greve G., Bachinski L., Friedman D.L., Czernuszewicz G., Anan R.,  
 RA Towbin J., Seidman C.E., Roberts R.;  
 RT "Isolation of a de novo mutant myocardial beta MHC protein in a  
 RT pedigree with hypertrophic cardiomyopathy.";  
 RL Hum. Mol. Genet. 3:2073-2075(1994).  
 RN [22]  
 RP VARIANTS CMH1 CYS-513; ARG-716 AND TRP-719.  
 RX MEDLINE=94110336; PubMed=882798;  
 RA Anan R., Greve G., Thierfelder L., Watkins H., McKenna W., Solomon S.,  
 RA Vecchio C., Shono H., Nakao S., Tanaka H., Mares A. Jr., Towbin J.A.,  
 RA Spirito P., Roberts R., Seidman J.G., Seidman C.E.;  
 RT "Prognostic implications of novel beta cardiac myosin heavy chain gene  
 RT mutations that cause familial hypertrophic cardiomyopathy.";  
 RL J. Clin. Invest. 93:280-285(1994).  
 RN [23]  
 RP VARIANT CMH1 THR-797.  
 RX MEDLINE=96047159; PubMed=7581410;  
 RA Moollan J.C., Brink P.A., Corfield V.A.;  
 RT "Identification of a novel Ala797Thr mutation in exon 21 of the beta-  
 RT myosin heavy chain gene in hypertrophic cardiomyopathy.";  
 RL Hum. Mutat. 6:197-198(1995).  
 RN [24]  
 RP VARIANT CMH1 CYS-453.  
 RX MEDLINE=96209901; PubMed=8655135;  
 RA Ko Y.-L., Chen J.-J., Tang T.-K., Cheng J.-J., Lin S.-Y., Liou Y.-C.,  
 RA Kuan P., Wu C.-W., Lien W.-P., Liew C.-C.;  
 RT "Malignant familial hypertrophic cardiomyopathy in a family with a  
 RT 453Arg->Cys mutation in the beta-myosin heavy chain gene:  
 RT coexistence of sudden death and end-stage heart failure.";  
 RL Hum. Genet. 97:585-590(1996).  
 RN [25]  
 RP VARIANTS CMH1 THR-349 AND TRP-719.  
 RX MEDLINE=98204402; PubMed=9544842;  
 RA Uesche B., Uhl K., Weist B., Schroder D., Meitinger T.,  
 RA Dohleman C., Vosberg H.-P.;  
 RT "A high risk phenotype of hypertrophic cardiomyopathy associated with

Query Match 6.8%; Score 163; DB 1; Length 1935;  
 Best Local Similarity 22.1%; Pred. No. 0.52;  
 Matches 64; Conservative 62; Mismatches 107; Indels 56; Gaps 9;

QY 31 LQCLIQSFETAPSRCTPCPOCRIOVGKRTIINKLFFDLAQBEEVNLDRF----- 78  
 DB 1143 LEEISLERLEENGAGTSQVIMNKKREAFPGKMRDL---DEATLQHEHTAATAALRKHADS 1199  
 QY 79 ----LKNELDNVRAQLSKQKKEKRDSDVITDTLRDTLEERNATVYSLQALGKEMLCSTL 135  
 DB 1200 VAEIGEIOIDIMQRYKQKLEKEKSEFKLELDVYSNMQ-----IKAKANLEKMCRTL 1252  
 QY 136 KKQKLYEQQOODETKQKEAGRLRSKMT-----MEQIELLLQSLP----- 178  
 DB 1253 EDQMEHRSKAETQKRVNDLTQORAKIQTENGELSQQLDEKELI--SOLTGRKLYTQQ 1311  
 QY 179 -----EVEEMIRDMGVGQSAVEQLAVYCVSLKREYENLEKARKASGEVADKLKDLFSS 232

```

Db 1312 LEDLRKRLSEEVAKNALAHALQASARHDCDLRQYEEETEARKEIOLRVLSKANSEVAQW 1371
Oy 233 RSKLOT-----VSELDQAKLELSAQKDLSAQKEMSLSKKLTMLOET 277
Db 1372 RTVETDAIQRTELEAKKRL-AQR-LQEAEEAVEAVNAKCSLEKT 1417

RESULT 37
MYSS_CHICK
ID MYSS_CHICK STANDARD: PRT; 1938 AA.
AC P13538; 013228;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, adult.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEGHORN; TISSUE=Pectoralis muscle;
RA Chao T.H., Bandman E., Moore L.;
RT "Cloning, nucleotide sequence and characterization of a full-length
RT cDNA encoding the myosin heavy chain from adult chicken pectoralis
RT major muscle.";
RN [2]
RP SUBMITTED (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-205.
RC TISSUE=Pectoralis muscle;
RA Hayashida M., Maeta T., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: I.
RT Sequence of the amino-terminal 23 kDa fragment.";
RL J. Biochem. 110:54-59(1991).
RN [3]
RP SEQUENCE OF 206-636.
RC TISSUE=Pectoralis muscle;
RA MEDLINE=92041768; PubMed=1939028;
RA Komine Y., Maeta T., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: II.
RT Sequence of the 50 kDa fragment of subfragment-1.";
RL J. Biochem. 110:60-67(1991).
RN [4]
RP SEQUENCE OF 637-837.
RC TISSUE=Pectoralis muscle;
RA MEDLINE=92041769; PubMed=1939029;
RA Maeta T., Miyashita T., Matsuzono K., Tanioka Y., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: III.
RT Sequence of the 22 kDa fragment and the alignment of the 23 kDa, 50
RT kDa, and 22 kDa fragments.";
RL J. Biochem. 110:68-74(1991).
RN [5]
RP SEQUENCE OF 838-1938.
RC TISSUE=Pectoralis muscle;
RA MEDLINE=92041770; PubMed=1939030;
RA Maeta T., Yajima E., Nagata S., Miyashita T., Nakayama S., Matsuda G.;
RT "The primary structure of skeletal muscle myosin heavy chain: IV.
RT Sequence of the rod, and the complete 1,938-residue sequence of the
RT heavy chain.";
RL J. Biochem. 110:75-87(1991).
RN [6]
RP PRELIMINARY SEQUENCE OF 1-808.
RA MEDLINE=97092420; PubMed=3467365;
RA Maeta T., Hayashida M., Tanioka Y., Komine Y., Matsuda G.;
RT "The primary structure of the myosin head.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:416-420(1987).
RN [7]
RP SEQUENCE OF 842-1270.
RA MEDLINE=90121764; PubMed=2610940;
RA Watanabe B.;

```

```

RT "Complete amino-acid sequence of subfragment-2 in adult chicken
RT skeletal muscle myosin.";
RL Biol. Chem. Hoppe-Seyler 370:1027-1034(1989).
RN [8]
RP SEQUENCE OF 852-1108.
RX MEDLINE=89374803; PubMed=2775482;
RA Watanabe B.;
RT "Amino-acid sequence of the short subfragment-2 in adult chicken
RT skeletal muscle myosin.";
RL Biol. Chem. Hoppe-Seyler 370:549-558(1989).
RN [9]
RP SEQUENCE OF 1145-1270.
RX MEDLINE=89228549; PubMed=2713098;
RA Watanabe B.;
RT "Amino-acid sequence of the hinge region in chicken myosin
RT subfragment-2.";
RL Biol. Chem. Hoppe-Seyler 370:55-61(1989).
RN [10]
RP SEQUENCE OF 1857-1938 FROM N.A.
RX MEDLINE=87217964; PubMed=3034534;
RA Moriarty D.M., Barringer K.J., Dodgson J.B., Richter H.E.,
RA Young R.B.;
RT "Genomic clones encoding chicken myosin heavy-chain genes.";
RL DNA 6:91-99(1987).
RN [11]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-843.
RX MEDLINE=93303624; PubMed=8316857;
RA Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R.,
RA Tomchick D.R., Benning M.M., Winkelman D.A., Wesenberg G.,
RA Holden H.M.;
RT "Three-dimensional structure of myosin subfragment-1: a molecular
RT motor.";
RL Science 261:50-58(1993).
RN [12]
RP FUNCTION: MUSCLE CONTRACTION. MYOSIN IS A PROTEIN THAT BINDS TO
RP F-ACTIN AND HAS ATPASE ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
RN [13]
RP SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
RP HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
RP AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
RN [14]
RP SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
RN [15]
RP DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
RP CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
RP CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
RN [16]
RP PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
RP ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
RN [17]
RP MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
RP MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
RP SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
RP SUBFRAGMENT (S2).
RN [18]
RP SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
RN [19]
RP SIMILARITY: CONTAINS 1 IQ DOMAIN.
RN [20]
RP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement. (See http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch.)
RN [21]
RP EMBL: U87231; AAB47555.1; -
RP EMBL: M16557; AAA48970.1; -
RP PIR: PX0050; PX0051.
RP PIR: A26821; A26821.
RP PIR: S02082; S02082.
RP PIR: S04501; S04501.
RP PIR: S05515; S05515.
RP PDB: 2MYS; 11-JAN-97.
RP InterPro: IPR004008; IO.
RP InterPro: IPR004009; Myosin_N.
RP InterPro: IPR002928; Myosin_Tail.
RP InterPro: IPR001609; myosin_head.
RP Pfam: PF00612; IO; 2.
RP Pfam: PF00663; myosin_head; 1.
DR

```

DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR Prodom; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IQ; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PS50096; IQ; 1.  
 KW Myosin; Muscic protein; Coiled coil; Thick filament; Actin-binding;  
 KW ATP-binding; Methylation; Alkylation; Phosphorylation; Acetylation;  
 KW Calmodulin-binding; Multigene family; 3D-structure.  
 FT INT\_MET 0  
 FT DOMAIN 1 782 MYOSIN HEAD-LIKE.  
 FT DOMAIN 783 812 IQ.  
 FT DOMAIN 838 840 HINGE.  
 FT DOMAIN 841 1938 COILED COIL (POTENTIAL).  
 FT NP\_BIND 179 186 ATP (POTENTIAL).  
 FT DOMAIN 657 679 ACTIN-BINDING.  
 FT DOMAIN 759 773 ACTIN-BINDING.  
 FT MOD\_RES 1 1 ACETYLATION.  
 FT MOD\_RES 35 35 METHYLATION (MONO-).  
 FT MOD\_RES 130 130 METHYLATION (TRI-).  
 FT MOD\_RES 551 551 METHYLATION (TRI-).  
 FT MOD\_RES 755 755 METHYLATION (MONO-).  
 FT MOD\_RES 697 697 ALKYLATION (SH-1).  
 FT MOD\_RES 707 707 ALKYLATION (SH-2).  
 FT CONFLICT 907 907 C -> Q (IN REF. 7 AND 8).  
 FT CONFLICT 980 980 L -> F (IN REF. 1).  
 FT CONFLICT 1343 1343 E -> D (IN REF. 5).  
 FT CONFLICT 1545 1545 S -> A (IN REF. 5).  
 FT CONFLICT 1796 1797 HV -> OL (IN REF. 5).  
 FT CONFLICT 1830 1830 S -> A (IN REF. 5).  
 FT CONFLICT 1863 1863 I -> V (IN REF. 10).  
 FT CONFLICT 1929 1931 IHG -> FH (IN REF. 10).  
 SQ SEQUENCE 1938 AA; 223013 MW; EDD01CEA2681E10F CRC64;

Query Match 6.88; Score 163; DB 1; Length 1938;  
 Best Local Similarly 22.3%; Pred. No. 0.52;  
 Matches 65; Conservative 58; Mismatches 107; Indels 62; Gaps 9;

QY 31 LOC105FETAPSRPTCPOCRIOVGKRTIINKLFFDLAQEENVDREF----- 78  
 DB 1145 LEEISERLEEGAGATAAIEENKKREAEFOAMRDL--EATLQHEHTAALAKKKHADQ 1201  
 QY 79 ---LKNELDVNRAQLSQKDEKRDQVYIDLPDLTEERNATVVSLOALGKAMLCSTL 135  
 DB 1202 TAELEGEDINDNQRYKOKLEKSKSLKMEIDLASNME-----SVKAKANLEKMCRTL 1254  
 QY 136 KKQKKYLEQOODE-----TKQAEAGRL-----RSKMKTMEQI 169  
 DB 1255 EDOLSEIKTEEQNQRMINDLTQARLQETGEGYSRQAEKDKALISLSKQGFQOI 1314  
 QY 170 ELLIQLSLPEVEEMIRMGVQGSVEQLAVYCVSLKKEYEMLKARKASGVADKLKLD 229  
 DB 1315 EELKR---HLEERIKKNALAHALQSARHDCCELLREOYEEOQAKGELORALSKANSEV 1370  
 QY 230 FSSRSKIQT---VYSELDAQLELKSQKDLQASDKREIMSKKRLMLQET 277  
 DB 1371 AQMTKRYETDAIQRTLEELERKKKL--AQR-LQDAEHEVAVNAKCSLEKT 1419  
 RESULT 38  
 KFA4\_HUMAN STANDARD; PRT; 1232 AA.  
 ID KFA4\_HUMAN 095239; 09NNT6; 09NY24;  
 AC 15-JUL-1999 (Rel. 38, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Chromosome-associated kinesin KIF4A (Chromokinesin).  
 GN KIF4A OR KIF4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;

RN [1]  
 RC SEQUENCE FROM N.A.  
 RP TISSUE=Lymphocytes;  
 RA Villard L.;  
 RL Submitted (MUG-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RC SEQUENCE FROM N.A.; AND TISSUE SPECIFICITY.  
 RP MEDLINE=20435301; PubMed=10978527;  
 RA Oh S.J., Hahn H., Torrey T.A., Shin H., Choi W., Lee Y.M.,  
 RA Morse H.C. III, Kim W.;  
 RT "Identification of the human homologue of mouse Kif4, a kinesin  
 superfamily motor protein.";  
 RL Biochim. Biophys. Acta 1493:219-224(2000).  
 RN [3]  
 RC TISSUE=Retinoblastoma;  
 RA Rentsch A., Neumann T., Rommerskirch W.;  
 RP Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RC SEQUENCE OF 128-1232 FROM N.A.  
 RP MEDLINE=97311419; PubMed=9168136;  
 RA Van R.-T., Wang S.-Z.;  
 RT "Increased chromokinesin immunoreactivity in retinoblastoma cells.";  
 RL Gene 189:263-267(1997).  
 CC -1- FUNCTION: REQUIRED FOR MITOTIC CHROMOSOMAL POSITIONING AND BIPOLAR  
 SPINDLE STABILIZATION (BY SIMILARITY)  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR, ASSOCIATED WITH MITOTIC  
 CHROMOSOMES (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEMATOPOIETIC TISSUES,  
 FETAL LIVER, SPLEEN, THYMUS AND ADULT THYMUS AND BONE MARROW.  
 CC LOWER LEVELS ARE FOUND IN HEART, TESTIS, KIDNEY, COLON AND LUNG.  
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.  
 CC CHROMOKINESIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF179308; AAD51855.1; -;  
 CC EMBL: AF071592; AAD05492.2; -;  
 CC EMBL: AJ271784; CAB75427.1; -;  
 CC EMBL: AF277375; AAF86334.1; -;  
 CC HSSP: P17119; 3KAR.  
 DR InterPro: IPR001752; kinesin.  
 DR PRINTS: PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC. 1.  
 DR PROSITE; PS00411; KINESIN\_MOTOR\_DOMAIN1; 1.  
 DR PROSITE; PS50067; KINESIN\_MOTOR\_DOMAIN2; 1.  
 KW Motor protein; Microtubules; ATP-binding; DNA-binding;  
 KW Nuclear protein; Coiled coil.  
 FT DOMAIN 1 349 KINESIN-MOTOR.  
 FT DOMAIN 350 999 COILED COIL (BY SIMILARITY).  
 FT DOMAIN 1000 1232 GLOBULAR.  
 FT NP\_BIND 88 95 ATP (POTENTIAL).  
 FT CONFLICT 223 223 R -> G (IN REF. 2).  
 FT CONFLICT 231 231 S -> T (IN REF. 4).  
 FT CONFLICT 286 286 V -> A (IN REF. 2).  
 FT CONFLICT 422 422 L -> W (IN REF. 2).  
 FT CONFLICT 564 564 L -> H (IN REF. 4).  
 FT CONFLICT 564 564 L -> P (IN REF. 2).  
 FT CONFLICT 600 600 K -> E (IN REF. 3).  
 FT CONFLICT 668 668 R -> K (IN REF. 3 AND 4).  
 FT CONFLICT 928 928 Q -> P (IN REF. 1).  
 FT CONFLICT 958 959 QL -> RL (IN REF. 3).  
 FT CONFLICT 960 960 L -> Q (IN REF. 1).  
 FT CONFLICT 996 997 LL -> S (IN REF. 4).  
 FT CONFLICT 1003 1014 OKHLPRDTLLSP -> RTLPRIPIFYLQ (IN REF. 4).



FT CONFLICT 1022 1022 P -> Q (IN REF. 2).  
 FT CONFLICT 1077 1077 K -> N (IN REF. 2).  
 FT CONFLICT 1138 1138 G -> S (IN REF. 2).  
 SQ SEQUENCE 1232 AA; 139908 MW; FF74052A17A6E8F7 CRC64;

Query Match 6.8%; Score 162.5; DB 1; Length 1232;  
 Best Local Similarity 21.4%; Pred. No. 0.33;  
 Matches 104; Conservative 67; Mismatches 143; Indels 173; Gaps 21;

QY 65 DLAEENVDREFLNK-----ELDNVRAQLSQKDKERDS-QVITDRLDTLEE 113  
 DB 777 DVALLKEK---KESGPNPKLRRTSLTEVRQVS-----ESDSITKQIESLETMEF 829  
 QY 114 RNATVYSLOALGKAEMLCTLKQKKYLLQOQDETQKQAE-AGRLRSKMKTKMEQELL 172  
 DB 830 RSKQIADLQOKLDAE-----SEDRPKRMENATITLAEACALKYLLGEL 874  
 QY 173 LOSQLPEVEEMIRDMVGQSGAVEQLAVYCVSLKEKENLKEARKASGEVADKLKRLDF-- 230  
 DB 875 VSKRI-QVSKL-----ESLQSKTSKADMK--MLFERNHFAFETELQAEIVRM 923  
 QY 231 --SSRSKLQTVSELDQAKTELKSAQKDLQADKEIMS----- 266  
 DB 924 EOHQOEKVLTLQLOOSQMAEKQLEESVSEKEQLLTLKCODELEKKREVCQEQQL 983  
 QY 267 -----LKKITLMQETLNLPRVASEVVDRLVLESPAPVEVNLKLRPRRDIDLNATF 320  
 DB 984 LRENEITKQTLTLQVASRQKHLEKDF-----LISP-----DSSF 1018  
 QY 321 DVDTPAPRSOSHGYE-----KLCEKSHSPIOD-----VPKIKCG 359  
 DB 1019 EYVPRKPRPRVKEKLEQGMIDIEDLKCYGSEHSVNEHEDGDDDEDEMKRTKLKYV 1078  
 QY 360 PKRESQSLGQSCAGPPEDELVGAPPIFYRNAILGQKQKPRPSSESSCKDVV----- 413  
 DB 1079 SRKNIQ-----GCSCKG-----WCGNKQCGCRKQKDCGDCGCCDPTKC 1117  
 QY 414 ---RTGFDGLG-----GRTEFIPTDPIV---MIRPLVPKPKR-----VPRQV 450  
 DB 1118 RNRQGGDSIGTYVERTDSGSKFLIEDPTTEVTEGSLFFNPVCPATNSKILKMCDBQVL 1177  
 QY 451 RVKTVPS 457  
 DB 1178 SKTKPPA 1184

RESULT 39  
 MYS\_AEOIR  
 ID MYS\_AEOIR STANDARD; PRT; 1938 AA.  
 AC P24733;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Myosin heavy chain, striated muscle.  
 OS Aequipecten irradians (Bay scallop).  
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Pectinoida;  
 OC Pectinoida; Pectinidae; Argopecten.  
 OX NCBI\_TaxID=31199;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adductor muscle;  
 RX MEDLINE=92011595; PubMed=1917970;  
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;  
 RT Complete primary structure of a scallop striated myosin heavy  
 chain. Sequence comparison with other heavy chains reveals regions  
 that might be critical for regulation.;  
 RL J. Biol. Chem. 266:18469-18476(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adductor muscle;  
 RX MEDLINE=91088319; PubMed=2263488;  
 RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;

RT "Nucleotide sequence of full length cDNA for a scallop striated  
 RT muscle myosin heavy chain.";  
 RL Nucleic Acids Res. 18:7158-7158(1990).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.  
 RX MEDLINE=94173332; PubMed=8127365;  
 RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,  
 RT Szent-Gyorgyi A.G., Cohen C.;  
 RT "Structure of the regulatory domain of scallop myosin at 2.8-A  
 RT resolution.";  
 RL Nature 368:306-312(1994).  
 RN [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.  
 RX MEDLINE=96419133; PubMed=8805510;  
 RA Houdusse A., Cohen C.;  
 RT "Structure of the regulatory domain of scallop myosin at 2-A  
 RT resolution: implications for regulation.";  
 RL Structure 4:21-32(1996).  
 CC CC  
 CC -1- FUNCTION: MUSCLE CONTRACTION.  
 CC -1- ACTIVITY THAT IS ACTIVATED BY F-ACTIN.  
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2  
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)  
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).  
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.  
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.  
 CC -----  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; X55714; CAA39247.1; -;  
 DR PIR; S13557; S13557.  
 DR PIR; A40997; A40997.  
 DR PDB; 1SCM; 30-APR-94.  
 DR PDB; 1WDC; 11-JUL-96.  
 DR InterPro: IPR000048; IO.  
 DR InterPro: IPR004009; Myosin\_N.  
 DR InterPro: IPR002928; Myosin\_tail.  
 DR InterPro: IPR002017; Spectrin.  
 DR InterPro: IPR001609; myosin\_head.  
 DR Pfam; PF00612; IO; 2.  
 DR Pfam; PF00063; myosin\_head; 1.  
 DR Pfam; PF02736; Myosin\_N; 1.  
 DR Pfam; PF01576; Myosin\_tail; 1.  
 DR PRINTS; PR00193; MYOSINHEAVY.  
 DR ProDom; PD000355; myosin\_head; 1.  
 DR SMART; SM00015; IO; 1.  
 DR SMART; SM00242; MYSC; 1.  
 DR PROSITE; PSS0096; IO; 1.  
 KW Myosin: Muscle protein: Coiled coil: Thick filament: Actin-binding;  
 KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.  
 KW DOMAIN 1 777  
 FT DOMAIN 1 778 805  
 FT DOMAIN 1 778 805  
 FT DOMAIN 1 778 805  
 FT NP\_BIND 176 183  
 FT NP\_BIND 176 183  
 FT MOD\_RES 693 693  
 FT MOD\_RES 693 693  
 FT MOD\_RES 693 693  
 FT HELIX 778 821  
 FT TURN 822 823  
 FT TURN 825 833  
 FT TURN 834 835  
 SQ SEQUENCE 1938 AA; 222821 MW; A5CCE4127D1A4896 CRC64;

Query Match 6.8%; Score 162.5; DB 1; Length 1938;  
 Best Local Similarity 25.1%; Pred. No. 0.55;



```

DR EMBL: AF091711; AAD39719.1; -.
DR EMBL: AB018346; BAA34523.1; -.
DR EMBL: AC000066; AAC60380.1; ALT_FRAME.
DR MTM: 604001; -.
KW Coiled coil; Alternative splicing; Polymorphism.
FT DOMAIN 2534 2567 PKA-RIT SUBUNIT BINDING DOMAIN.
FT DOMAIN 164 914 COILED COIL (POTENTIAL).
FT DOMAIN 944 1022 COILED COIL (POTENTIAL).
FT DOMAIN 1100 1185 COILED COIL (POTENTIAL).
FT DOMAIN 1253 1280 COILED COIL (POTENTIAL).
FT DOMAIN 1336 1392 COILED COIL (POTENTIAL).
FT DOMAIN 1434 1459 COILED COIL (POTENTIAL).
FT DOMAIN 1585 1659 COILED COIL (POTENTIAL).
FT DOMAIN 1857 2455 COILED COIL (POTENTIAL).
FT DOMAIN 2544 2561 COILED COIL (POTENTIAL).
FT DOMAIN 2603 2776 COILED COIL (POTENTIAL).
FT DOMAIN 3065 3092 COILED COIL (POTENTIAL).
FT DOMAIN 3124 3470 COILED COIL (POTENTIAL).
FT DOMAIN 3587 3689 COILED COIL (POTENTIAL).
FT DOMAIN 3726 3730 COILED COIL (POTENTIAL).
FT DOMAIN 203 292 POLY-LEU.
FT DOMAIN 321 1010 GLN-RICH.
FT DOMAIN 1846 2772 GLU-RICH.
FT DOMAIN 17 28 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT VARSPPLIC 1637 1642 OLQEEI -> LATRD (IN ISOFORM 4).
FT VARSPPLIC 1643 3911 MISSING (IN ISOFORM 4).
FT VARSPPLIC 2175 2182 MISSING (IN ISOFORM 3).
FT VARSPPLIC 2175 2183 SADTFORVE -> Q (IN ISOFORM 6).
FT VARSPPLIC 2895 2907 VEGFYNNCESTLC -> GSSIPPLAHSADAYOTREISS
  (IN ISOFORM 2, ISOFORM 3 AND ISOFORM 6).
FT VARSPPLIC 2895 2948 MISSING (IN ISOFORM 5).
FT VARSPPLIC 2901 3911 STIOFHAGMR -> ALSLTTSWQHHSAPRTAPLFEILSH
  SLG (IN ISOFORM 6).
FT VARIANT 1347 1347 K -> KO.
  /FTID=VAR_010926.
FT CONFLICT 76 76 E -> Q (IN REF. 3).
FT CONFLICT 475 475 M -> I (IN REF. 3).
FT CONFLICT 554 554 R -> G (IN REF. 3).
FT CONFLICT 638 638 R -> S (IN REF. 3).
FT CONFLICT 663 663 N -> S (IN REF. 3).
FT CONFLICT 913 913 H -> N (IN REF. 3).
FT CONFLICT 956 956 K -> N (IN REF. 3).
FT CONFLICT 980 982 OKH -> PKP (IN REF. 1 AND 2).
FT CONFLICT 997 997 Q -> P (IN REF. 1 AND 2).
FT CONFLICT 1001 1001 Q -> P (IN REF. 1 AND 2).
FT CONFLICT 1020 1020 N -> D (IN REF. 3).
FT CONFLICT 1028 1028 V -> E (IN REF. 3).
FT CONFLICT 1626 1626 R -> P (IN REF. 3).
FT CONFLICT 1703 1703 N -> T (IN REF. 3).
FT CONFLICT 1707 1707 V -> G (IN REF. 3).
FT CONFLICT 1802 1803 MISSING (IN REF. 5).
FT CONFLICT 1843 1843 A -> P (IN REF. 3).
FT CONFLICT 1956 1956 I -> V (IN REF. 3).
FT CONFLICT 2027 2027 V -> D (IN REF. 5).
FT CONFLICT 2157 2158 EI -> HE (IN REF. 7).
FT CONFLICT 2169 2169 E -> V (IN REF. 3).
FT CONFLICT 2514 2514 L -> R (IN REF. 3).
FT CONFLICT 2851 2851 I -> N (IN REF. 8).
FT CONFLICT 2957 2957 E -> D (IN REF. 3).
FT CONFLICT 2983 2983 P -> S (IN REF. 3, 7 AND 8).
FT CONFLICT 3087 3087 Q -> H (IN REF. 3).
FT CONFLICT 3218 3218 Q -> H (IN REF. 3).
FT CONFLICT 3307 3309 ESE -> OSO (IN REF. 3).
FT CONFLICT 3751 3751 P -> A (IN REF. 3).
FT CONFLICT 3833 3833 T -> S (IN REF. 3).
SQ SEQUENCE 3911 AA; 453664 MM; 3FB1CB1C819B47AA CRC64;

```

```

Query Match 6.8%; Score 162.5; DB 1; Length 3911;
Best Local Similarity 22.2%; Pred. No. 1.2;
Matches 96; Conservative 80; Mismatches 176; Indels 81; Gaps 18;
QY 41 APSRTCPQCRIOVGKRTINKLFFDL-----AQEEENVLDLR--EFLKNE 82

```

```

Db 1940 ADEKTLFEHQIQ-EKTDIIDRLQEQELLCASNRQLQELAEQOQIQEERRELLSRQKEAMKAE 1998
QY 83 LDVVRQALSKD-----KEKDSOVIITDLRDTL-----BERNA 116
Db 1999 AGPEEQDLQETKRLMKKELEVOCQAEKVRDLDQVKALEIDVEROVSRFTIELDEBKYN 2058
QY 117 TVVSLQOALGKAEMLCSTLKKOMKYLEQO----QDETQKQAEAGRLRSKMTMEQIELL 172
Db 2059 ELMDLRO---QNALQKQLEKMKRFLDEQALIDREHERDVQOETQKLEQDLKVVPRFQPI 2115
QY 173 LOSQLPEVEEMIRDMGVGQSAVEQOLAVYCVSLKKE-YENLKEARKASGEVADRLRKDFPS 231
Db 2116 SEHQTRVEQDLAHMLKKTCKSCSELLSKRQQLQRDQGERNEELKLEFRVRE-LEQALLV 2174
QY 232 SRSKLQTVSELD---QAKLELK-----SAQKD-LOSADKETMSLKKTLMLQETLNP 281
Db 2175 SADTFORVEDRKRHFGAWEAKPELSLEVQLOAERDAIDRKEKEITNLLEEQLEQFREBLE-- 2232
QY 282 PVASETVDRVLV-----ESPAPVEVNLKLRPSRFDIDLNATFDVDPAPRPSSSH 334
Db 2233 -NKNEEVQQLHMQLEIQKKESTTRIQ-ELQENKLFKDDME-KLGALIKESDAMSTODH 2289
QY 335 GYEEKLC-LEKSHSPIQDVPKRIKCGPRKESQSLGQSCAGEPDELYGAFPIFVRNA 392
Db 2290 VLGKFAQIIQEKKEVEIDQJNEQVTK---LQQLKITTDKVIIEKKNELRIDLETQIECL 2346
QY 393 ILGOKQPKRPRSE 405
Db 2347 MSDQECVKNRREE 2359

```

Search completed: September 4, 2002, 16:17:19  
Job time: 478 sec

